In situ genome sequencing resolves DNA sequence and structure in intact biological samples

Andrew C. Payne[1,2]*, Zachary D. Chiang[2,3]*, Paul L. Reginato[1,2,4,5,6]*, Sarah M. Mangiameli[2], Evan M. Murray[2], Chun-Chen Yao[2,7], Styliani Markoulaki[8], Andrew S. Earl[3], Ajay S. Labade[3], Rudolf Jaenisch[8,9], George M. Church[5,6], Edward S. Boyden[1,4,7,10,11,12,13]†‡, Jason D. Buenrostro[2,3]†‡, Fei Chen[2,3]†‡

[1]Media Arts and Sciences, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA. [2]Broad Institute of MIT and Harvard University, Cambridge, MA 02142, USA. [3]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA. [4]Department of Biological Engineering, MIT, Cambridge, MA, 02139, USA. [5]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. [6]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115, USA. [7]Department of Brain and Cognitive Sciences, MIT, Cambridge, MA 02139, USA. [8]Whitehead Institute for Biomedical Research, Cambridge, MA 02139, USA. [9]Department of Biology, MIT, Cambridge, MA 02139, USA. [10]McGovern Institute, MIT, Cambridge, MA 02139, USA. [11]Koch Institute, MIT, Cambridge, MA 02139, USA. [12]Howard Hughes Medical Institute, Cambridge, MA 02139, USA. [13]Centers for Neurobiological Engineering and Extreme Bionics, MIT, Cambridge, MA 02139, USA.

*These authors contributed equally to this work.
†These authors contributed equally to this work.
‡Corresponding author. Email: edboyden@mit.edu (E.S.B.); jason_buenrostro@harvard.edu (J.D.B.); chenf@broadinstitute.org (F.C.)

Understanding genome organization requires integration of DNA sequence and 3D spatial context, however, existing genome-wide methods lack either base-pair sequence resolution or direct spatial localization. Here, we describe in situ genome sequencing (IGS), a method for simultaneously sequencing and imaging genomes within intact biological samples. We applied IGS to human fibroblasts and early mouse embryos, spatially localizing thousands of genomic loci in individual nuclei. Using these data, we characterized parent-specific changes in genome structure across embryonic stages, revealed single-cell chromatin domains in zygotes, and uncovered epigenetic memory of global chromosome positioning within individual embryos. These results demonstrate how in situ genome sequencing can directly connect sequence and structure across length scales from single base pairs to whole organisms.

The genome of an organism encodes not only its genes, but also principles of spatial organization that regulate gene expression and control cellular function (*1*, *2*). Accordingly, mapping spatial genome organization at high resolution is important for understanding its diverse regulatory roles in health, disease, and development (*3*, *4*). Principles of genome architecture have mostly been uncovered by methods based on DNA sequencing of chromatin contacts (*5*), such as Hi-C (*6*), and methods which probe targeted genomic loci using microscopy, such as DNA fluorescence in situ hybridization (FISH) (*7*). Hi-C applied to populations of cells has revealed genome-wide organizing principles (*8*–*12*), and single-cell variations have uncovered cell-to-cell heterogeneity (*13*–*17*). DNA FISH has similarly revealed genome architecture at single-cell resolution (*18*, *19*). More recent studies have shown how these approaches can complement each other by imaging Hi-C defined features in single cells, characterizing their heterogeneity, and validating inferred differences in chromatin conformation within and across cell types (*20*–*26*) However, these methods cannot currently be applied jointly on the same cell, and a method to simultaneously sequence and image genomes in single cells is lacking. Efforts which combine Hi-C with microscopy (*16*, *27*), or efforts which make FISH more like sequencing via single-nucleotide

polymorphism (SNP) specific probes (*23*, *28*), have broken important conceptual ground, but they remain limited in their imaging or sequencing throughput. Accordingly, questions requiring both genomic and spatial analysis in single cells have been difficult to address.

In situ genome sequencing workflow

Here we present a method for in situ genome sequencing (IGS). IGS enables DNA sequencing directly within intact biological samples, spatially localizing genome-wide paired-end sequences in their endogenous context and thus bridging sequencing and imaging modalities for mapping genomes. Our in situ sequencing workflow introduces innovations in three phases: in situ library construction, multimodal sequencing of libraries, and computational integration of spatial and genetic information.

In the first phase, we create an in situ sequencing library within fixed samples by amplifying an untargeted sampling of the genome in its native spatial context. To do this, we fix and treat samples using methods optimized for DNA FISH (*28*, *29*). Next, we use Tn5 transposase to randomly incorporate DNA sequencing adaptors into fixed genomic DNA by in situ transposition, preserving genomic fragments in their native spatial positions (*30*). We circularize these fragments in

situ by ligation of two DNA hairpins containing a unique molecular identifier (UMI) and primer sites used for subsequent multi-modal DNA sequencing (Fig. 1A, ii and iii). We then clonally amplify the resulting circular templates by rolling circle amplification, yielding in situ DNA sequencing libraries with up to thousands of spatially localized amplicons per nucleus (fig. S1). We also developed a method for modulating the effective density of sequencing libraries to optimize the number of resolvable amplicons (fig. S2). Together, this provides an approach to clonally amplify untargeted samples of a genome, creating approximately 400-500 nm sized features for in situ sequencing (fig. S3).

In the second phase of our workflow, we sought to use reported (31) in situ sequencing protocols to determine the sequence and 3D positions of amplicons. However, current in situ sequencing methods yield short single-end reads (at most 30 bases), and are limited by imaging time (31). This poses a challenge for genome sequencing: the human genome encodes 3 billion bases and includes highly repetitive regions, requiring long paired-end sequencing reads to resolve many regions of the genome. To address this challenge, we combined in situ sequencing with high-throughput paired-end DNA sequencing. To do this, we first read amplicon-specific UMIs within fixed samples using sequential rounds of in situ sequencing by ligation (SBL) and fluorescence imaging (31) (Fig. 1Bi and movie S1). Immunostaining followed by additional cycles of imaging may also be performed following in situ sequencing. We then dissociate the in situ amplicons and amplify them using PCR to produce an in vitro sequencing library (Fig. 1B, ii and iii), which we sequence on a conventional Illumina sequencer (henceforth referred to as ex situ sequencing) to obtain 150 bp paired-end genomic reads tagged with in situ sequenced, spatially resolved UMIs. This multimodal sequencing strategy allows us to preserve spatial information while leveraging the accuracy and read-length of paired-end sequencing on the Illumina platform, which is crucial for aligning individual reads to millions of unique genomic loci.

In the third phase, we computationally match ex situ paired-end sequencing reads to in situ amplicon positions. Briefly, we deconvolve, register, and normalize fluorescence images in order to resolve the 3D centers of amplicons across multiple rounds of imaging [fig. S4 (29)]. We then quantify the fluorescence signal of each UMI-associated amplicon across four color channels over all rounds of in situ sequencing (Fig. 1Bi and fig. S5). The ex situ sequenced reads are next associated with spatial coordinates within nuclei through error-robust matching of in situ and ex situ sequenced UMIs (Fig. 1C). To do this, we implement a probabilistic matching approach using principles from single-bit error correction (32) (fig. S5). Collectively, the integration of these methods, which include developments across library construction,

sequencing, and computational analyses, enable IGS as a general strategy for spatially mapping paired-end reads (29).

Here, we apply IGS to 106 human fibroblasts (PGP1f) and 113 cells across 57 intact early mouse embryos at the PN4 zygote (n = 24), late 2-cell (n = 20), and early 4-cell (n = 13) stages of development (Fig. 1D). Across both experiments, 66.35% of clearly resolvable amplicons (87.6% in PGP1f, 61.0% in mouse embryos) were confidently matched to an ex situ genomic read (fig. S6). After cell filtering based on yield, karyotype, developmental stage, and cell cycle (29), this yielded a total of 286,335 spatially localized genomic reads (36,602 in PGP1f, table S1; 249,733 in mouse embryos, table S2) with a UMI-matching false discovery rate of 0.26% [1.70% in PGP1f, 0.05% in mouse embryo (29)]. Mapped amplicons scaled with nuclear volume, spanning a median of 328 ± 114 reads per nucleus (± SD) in the PGP1f cells, to a median of 3,909 ± 2,116, 2,357 ± 1,063, and 1,074 ± 622 reads per nucleus in zygote, 2-cell, and 4-cell stage embryos, respectively (fig. S7). Sequencing coverage across the hg38 and mm10 reference genomes was comparable to whole genome sequencing (fig. S8), and genomic reads did not show bias based on radial position (fig. S9) or chromatin accessibility (fig. S10). We also quantified the rate of detection for each genomic region across individual cells, as well as the distribution of genomic distance between sampled loci on the same chromosome (fig. S11).

For downstream analyses, we annotated each read based on spatial features such as inclusion in chromosome territories and distance to nuclear landmarks (nuclear lamina, centromeres, and nucleolar precursor bodies), as well as published genomic data including A/B compartments, lamina-associated domains, and GC content (table S3). The full embryo dataset can be interactively visualized at https://buenrostrolab.shinyapps.io/insituseq/.

Validation of in situ genome sequencing in cultured cells

To validate that our method detects features of spatial genome organization, we first examined the locations of chromosomes in interphase human male PGP1 fibroblasts (Fig. 2, A and B, and movie S2). We found that autosomal reads displayed a strong tendency to spatially colocalize into two distinct spatial regions, while allosomal reads were restricted to one region, confirming the known organization of chromosomes into territories (18) (fig. S12 and movie S3). To systematically define these territories, we used a maximum likelihood estimation approach to assign reads to homologous chromosome clusters using both the spatial and genomic positions of each read (fig. S13). For chromosomes with two spatially resolved homolog clusters, we found, via density-based thresholding, that 6.83% of reads did not spatially colocalize with either cluster (fig. S13), a larger fraction than our estimated 1.70% UMI-matching false discovery rate in PGP1f, which may be associated with long-range

chromosome looping (33). Following spatial clustering, we visualized genome-wide conformations of individual chromosomes in single diploid cells by connecting the reads in each cluster according to genomic position (Fig. 2C).

We proceeded to characterize the positions of diploid chromosome territories across single cells by calculating the average spatial distance between pairwise genomic locations across the genome, at 10 Mb resolution. We found that blocks of short intrachromosomal distances were strongly delineated along the diagonal of this pairwise distance matrix (Fig. 2D), consistent with genomic organization into chromosome territories described above. The matrix also shows enrichment of shorter pairwise distances between smaller chromosomes. Additionally, we observed a positive association between chromosome size and radial distance from the nuclear center (Fig. 2E). These observations indicate that small chromosomes tend to be in closer proximity near the nuclear center, consistent with prior studies in human fibroblasts (34). These results illustrate the ability of IGS to resolve diploid chromosome territories within the nuclei of single cells and to investigate the spatial positioning of chromosomes at scale.

Repetitive DNA elements, such as transposons and endogenous retroviruses, make up approximately 50% of the human genome (35, 36), and their localization is known to play a role in normal (37) and disease-associated (38) genome organization. Although FISH-based methods can measure the localization of targeted classes of repetitive sequences (34, 39), current approaches have not simultaneously mapped the localization of many classes of repetitive sequences across the nucleus. We applied IGS to simultaneously measure the localization of repetitive sequences across the genome. We focused on the 13.9% of spatially resolved reads that do not uniquely align to the reference genome (hg38), and aggregated them into ~250 classes of repetitive elements using Repbase (40) (fig. S14A). We found that the number of reads associated with each element was proportional to its observed frequency in hg38 (fig. S14B), enabling an unbiased approach to studying localization of repetitive elements.

Given our observations of radial patterns of chromosome positioning and the known radial organization of heterochromatin (41), we investigated whether repetitive elements displayed radial patterns. To do this, we compared the radial distribution of reads containing repetitive elements to permuted distributions to identify classes of repetitive elements with the strongest radial bias [Fig. 2F (29)]. We confirmed reports that Alu elements are depleted ~1 μm from the nuclear edge (34), further validating our approach (Fig. 2G). We also found that certain types of repetitive heterochromatin, such as satellite DNA, show enrichment toward the nuclear center, while others, including AT-rich L1 elements, are overrepresented at the nuclear edge. These findings

demonstrate the ability of IGS to simultaneously identify spatial localization patterns of many different repetitive sequences de novo in an untargeted and genome-wide manner. Having shown that IGS confirms known features of global genome organization, we next asked whether we could characterize the structures of individual chromosomes. DNA FISH and Hi-C studies have found that chromosome arms can be individually compartmentalized in a fashion similar to chromosome territories (44, 45). Independent localization of chromosome arms was apparent as stripes in the genome-wide distance matrix (Fig. 2D) and could be visualized in single chromosomes colored by their p and q arms (Fig. 2H). We computed pairwise distances for reads within each chromosome territory (Chrs 1-11 and Chr X; fig. S15) and fit a power law to this relationship as described (20). Separate treatment of p and q arms resulted in an improved fit compared to fitting of all of Chr 1 (fig. S16), in line with the expectation of differential scaling across the centromere. Indeed, intra-arm and inter-arm pairs of loci exhibited two scaling regimes when treated separately (Fig. 2I and fig. S17). In chromosomes where we had high coverage (Chr 1-11 and Chr X), inter-arm distances were significantly larger than intra-arm distance for the range of genomic distances present in both distributions (56-87 Mb for Chr 1, Fig. 2J; boxed region, Fig. 2I; K-S test, p < 10^{-16}, all chromosomes, fig. S18). These results extend observations of spatially polarized chromosome arms (42) and demonstrate the ability of IGS to characterize sub-chromosomal spatial structure. Taken together, these findings highlight the unique ability of IGS to simultaneously interrogate broad features of genome organization, including chromosome positioning, chromosome folding, and the localization patterns of repetitive sequences.

In situ genome sequencing in intact early mouse embryos
The spatial organization of the genome is extensively remodeled in early embryogenesis, as the initially separate parental genomes undergo major reorganization after fertilization to prime the organism for zygotic genome activation (ZGA) (43) and, subsequently, lineage-specific cell fates (44). Studies have linked chromatin and epigenetic remodeling to various phenomena including sequence-specific localization of chromatin to nuclear landmarks (45, 46), parent specific-chromatin domain organization in single cells (47–50), and developmental specification of clonal lineages (51–53). Given the importance of spatial features, sequence-specificity, and intercellular relationships in these phenomena, we sought to apply IGS in intact early embryos to characterize genome organization in early embryogenesis across length scales. We applied IGS to intact early mouse embryos (B6C3F1 females x B6D2F1 males) spanning the PN4 zygote (3,909 ± 2,116 reads/cell, median ± SD), late 2-cell (2,357 ± 1,063 reads/cell), and early 4-cell (1,074 ± 622 reads/cell) stages of development (Fig. 3A and movie S4).

Collectively, imaging and sequencing methods have shown that some of the structural changes in early development are associated with nuclear landmarks, such as the centromeres (54), nuclear lamina (46) and nucleolus precursor bodies (NPBs) (55). To demonstrate that IGS is compatible with other imaging modalities and to investigate the organizational roles of these landmarks, we performed co-immunostaining for CENP-A (centromere) and Lamin-B1 (nuclear lamina), in addition to staining with DAPI (used to locate NPBs). The resulting images were segmented and registered to the in situ sequencing data in 3D, enabling us to directly localize genomic reads relative to these landmarks [Fig. 3B (29)].

In order to resolve the maternal and paternal genomes within single cells, we first confirmed the presence of chromosome territories in all stages. As with the PGP1 fibroblasts, we found that reads originating from a particular autosome could generally be separated into two distinct spatial clusters per nucleus (or one cluster per allosome in male embryos) (movie S5). We then filtered cells based on yield, karyotype, developmental stage, and cell cycle [fig. S19 (29)]. After filtering, we found a nearly equivalent rate of reads that did not colocalize with chromosome territories as in PGP1f (6.95%). Relative to reads within territories, these non-colocalizing reads were significantly depleted from regions proximal to the nuclear lamina and NPBs (K-S test, p < 10^{-51} and 10^{-109} respectively; fig. S20).

To assign parent-of origin to each territory, we identified spatially localized reads that overlapped a genomic position with a heterozygous SNP in either of the parental strains. 1.40% and 1.64% of genomic reads were uniquely assigned to the maternal (B6C3F1) and paternal (B6D2F1) genomes respectively, resulting in an average of 67 haplotype-informative reads per cell. To validate these assignments, we visualized the positions of haplotype-informative reads in the PN4 zygote (Fig. 3C). At this stage of development, the parental genomes remain segregated in the larger paternal and smaller maternal pronuclei. Based on this known feature, we assigned each chromosome territory to either the maternal or paternal genome in a semi-supervised manner (Fig. 3D). We found that 97.1% of haplotype-informative reads were concordant with this assignment, where non-concordant reads may be attributable to genomic sequencing errors, UMI matching errors, or strain impurities. We then used our haplotype-informative reads to assign entire chromosome territories to parent-of-origin across the 2-cell and 4-cell stage embryos (29). This approach enables a strong majority of reads (82.26%), even those not overlapping a SNP, to be assigned to parent-of-origin through co-localization with haplotype-resolved reads in the same territory.

Developmental transitions in embryonic genome organization

We then sought to examine previously described principles of global genome organization, focusing on parental haplotype (56, 57), centromere-telomere position (16, 54), and GC content (6, 16) (Fig. 4A).

We began by examining the spatial separation of parental genomes, as imaging studies have shown that maternal and paternal chromatin are spatially polarized in early embryos (56, 57). To quantify the spatial separation of parental genomes across developmental stages, we analyzed the spatial inter-chromosomal neighbors of each read and calculated a haplotype spatial separation score (Fig. 4B). We then averaged the separation scores for all reads in each cell. We found that the mean separation score significantly decreased between the zygote and 2-cell stages and between the 2-cell and 4-cell stages (K-S test, p < 10^{-8} and p < 10^{-3}; Fig. 4C), consistent with earlier studies (56, 57). The standard deviation of mean separation scores increased with each developmental stage (SD = 0.015 for zygotes, 0.026 for 2-cell, 0.045 for 4-cell), indicating that the degree of parental genome intermixing is heterogeneous within the embryo. Further, we observed no evidence that particular chromosomes were more likely to break this separation than others (fig. S21). These results are concordant with the hypothesis that gradual mixing is a consequence of global chromosome repositioning following mitosis.

We next examined global spatial organization of the genome along the centromere-telomere axes of chromosomes. Mouse chromosomes are acrocentric and are known to be arranged in a Rabl-like configuration in early embryos, in which centromeres cluster toward one side of the nucleus and distal telomeres cluster toward the other (54, 56). To confirm this configuration in our data, we first measured the polarity of the CENP-A stain and found that centromeres in the 2 and 4-cell stages were significantly clustered toward one side of the nucleus (K-S test, p < 10^{-4} and p < 10^{-8}). To analyze this configuration for all chromosome positions, we assigned each read a centromere-telomere score based on its genomic position along its chromosome. When we visualized these scores in a nucleus from a 2-cell embryo, we observed that the centromere-telomere scores were highly polarized, which was supported by co-localization of the CENP-A immunostain (Fig. 4D). To quantify this polarization across all stages, we calculated a spatial neighborhood centromere-telomere score for each read (29). We then examined the relationship between centromere-telomere scores and neighborhood scores across all reads and observed much stronger correlation in the 2-cell and 4-cell stages (Pearson's r = 0.519 and 0.502) than in the zygote (r = 0.074; Fig. 4E). The functional consequences of this transition to a Rabl-like configuration in 2- and 4-cell embryos remains unclear. In other contexts, this configuration is thought to be an extension of anaphase chromosome positioning into interphase, perhaps without cellular function (58). On the other hand, it has also been

hypothesized to restrict chromatin entanglement (*59*); thus, it may be involved in constraining genome structure to enable the short cell cycles of the early embryo.

Finally, we examined the role of GC content in genome structure, which is strongly associated with A/B compartmentalization (*6*). To study this effect, we first visualized individual homologs of Chr 12 from zygotic pronuclei (Fig. 4F). We observed that genomic reads from GC-poor regions tended to localize to the periphery of the nucleus, in line with reports describing the localization of the inactive B compartment (*16*). To quantify this effect, we measured the correlation between GC content and distance to nuclear lamina across Chr 12 in all zygotes and observed that these two factors were correlated in both the paternal and maternal homologs (Spearman's ρ = 0.794, 0.649 respectively; Fig. 4G). We applied this approach to all chromosomes and found that the paternal homologs were significantly more correlated than their maternal counterparts in the zygote, but not in the 2-cell stage (K-S test, $p < 10^{-5}$, n.s.; Fig. 4H). This suggests that the degree of GC compartmentalization may be influenced by the differing biological histories of the pronuclei (*47*). Intriguingly, when we examined the relationship between genomic position and distance to nuclear landmarks, we observed that Chr X seemed to localize especially far from the nuclear lamina and toward the NPBs in paternal pronuclei (fig. S22). This finding extends models of the role of NPBs in establishing epigenetic asymmetry between parental X chromosomes rapidly after fertilization and in advance of imprinted X inactivation (*45*). Taken together, these results demonstrate the ability of IGS to characterize 3D genome organization across diverse developmental stages with parent-specific resolution, and with respect to nuclear landmarks.

Detection of single-cell domains chromatin domains in zygotes
Next, we used IGS to examine subchromosomal spatial organization. We focused on our data with the highest genomic resolution, the zygotic pronuclei, where large-scale parent-specific reorganization of chromatin is thought to play an important role in ZGA (*60*). First, we characterized the scaling relationship between mean spatial and genomic pairwise distance in the zygotic parental genomes. Consistent with previous reports (*47*), we found that each parental genome had distinct scaling properties (Fig. 5, A and B).

Reports analyzing genome structure in zygotes have suggested that paternal chromatin exhibits unusually weak higher order structure (> 2 Mb) (*47–49*). In accordance with these reports, we found that spatial distance matrices generated from the population ensemble indeed exhibited little off-diagonal structure (Fig. 5C and fig. S23, all chromosomes). Furthermore, evidence suggests that paternal zygotic chromatin exhibits unusually weak A/B compartmentalization

(*47–49*), and unusually large lamina associated domains (*46*). Our ensemble data corroborated these reports and correlated well with Hi-C (mean Pearson's r = 0.84) when analyzed in terms of lamin proximity (Fig. 5D and fig. S24, all chromosomes). However, when we examined single-cell distance matrices, we found that, unlike the ensemble, single paternal pronuclei generally exhibited large blocks of spatially associated chromatin (Fig. 5E, left). To distinguish these blocks from population-defined topological domains identified in Hi-C studies, we term them single-cell domains (SCDs).

To better understand the nature of SCDs in paternal zygotes, we systematically identified individual domains in single cells (*29*). The SCDs we identified corresponded well to spatially distinct clusters identified by visual inspection (Fig. 5E, right). When we examined SCDs across cells and chromosomes, we observed that they were large (median size 17.5 Mb, 10 Mb inter-quartile range) relative to canonical features defined by Hi-C, and had heterogeneous sizes and boundary positions (fig. S25). We proceeded to assess the strength of all SCD boundaries in single cells, and found they were significantly stronger and more variable than boundaries identified by the same method in the ensemble matrices [K-S test, $p < 10^{-17}$; 95% CI for Cohen's d (0.56, 0.77) determined by bootstrapping (*29*); Fig. 5F]. Together, these observations suggest that the weak ensemble structure may be explained by the variability of single-cell structures. Finally, we investigated the association of SCDs with nuclear landmarks, which may suggest organizing principles, and found that their boundaries and interiors were, respectively, significantly more lamin distal and more lamin proximal than expected [Fig. 5G (*29*)]. We found this striking in light of previous electron microscopy studies showing discrete micron-scale lamina-associated chromatin domains following ZGA (*61*).
Recently, both polymer simulations (*47*) and direct observation of chromatin structure (*22*) have shown how variable domain-like structures can exist in single cells when higher order ensemble structure is lacking. We speculate that, given the weak ensemble structure, the SCDs observed here may involve a similar phenomenon. The SCDs described here are larger than canonical Hi-C defined features and were not detected in earlier single-cell Hi-C studies in zygotes (*47*, *50*), perhaps because they are organized on length scales which are relatively less accessible to Hi-C measurements. It may be interesting to investigate the extent to which SCDs are governed by mechanisms related to the nuclear lamina, which perhaps modulates underlying epigenetic (*41*) or polymer-intrinsic (*47*) domain-forming behaviors of chromatin.

Epigenetic memory of global chromosome positioning within early embryos
Embryonic development is thought to involve epigenetic transmission of structural and regulatory features of

chromatin organization through clonal cell lineages within individual embryos (*53*). These mechanisms play an important role in breaking initial symmetry (*51*), engaging clonal lineage-specific gene expression programs (*62*), and cell fate commitment (*44*). Intercellular asymmetries influencing the developmental fate of clonal lineages have been reported as early as the 4-cell stage within individual embryos (*51*, *52*, *62*, *63*).

In order to study clonal lineage-specific features at the single-cell level, it is necessary to resolve and compare cells within the same embryo. Chromosome territories have been found to form early in interphase and subsequently maintain their relative positions until prophase (*18*, *64*, *65*), so we reasoned that comparison of chromosome positioning would be a robust way to quantify the similarity of global genome organization between interphase cells. Live-cell studies using non-specific photopatterning of the nucleus have demonstrated similarity in global genome organization between sister cells in culture (*65*), and indeed, visual inspection of pairs of chromosomes suggested that cells within an embryo share similar chromosomal positions (Fig. 6A). We quantified similarity by comparing single-cell autosome distance matrices of pairs of cells within and between individual embryos [Fig. 6, A and B (*29*)]. In 2-cell embryos, we found that global chromosome positioning in sister cells was significantly more correlated than in pairs of cells from different embryos (K-S test, p < 10^{-15}; Fig. 6C and fig. S26). These results suggest that cells within 2-cell embryos may share memory of their common initial chromosome positioning during zygotic metaphase, if not earlier.

Next, we asked whether the similarity shared by sister cells in 2-cell embryos might be epigenetically transmitted across the second cell division, i.e., to cousin cells. While earlier work has not found heritability in the radial positioning of individual loci (*66*, *67*), widely varying degrees of similarity in global genome organization between mother and daughter cells have been reported (*64*, *65*). We constructed putative clonal lineage trees within each 4-cell embryo by using the ranked correlations of autosome distance matrices for each pair of cells to classify putative sister and cousin cells. The most correlated pair of cells in each embryo was designated as one set of putative sister cells, thus implying the remainder of the tree [Fig. 6D and fig. S27 (*29*)]. As expected by the definition of the tree, we found that global chromosome positioning was significantly more correlated between putative sister cells than between pairs of cells from different 4-cell embryos. However, we also found that the same held true for putative cousin cells, which was not expected [K-S test, p < 10^{-14} and p < 10^{-3}; 95% CI for Cohen's d (0.55, 1.29) determined by bootstrapping (*29*); Fig. 6E and fig. S26]. Together, these results demonstrate clonal lineage-specific similarity in global chromosome positioning in early embryos, and imply that epigenetic memory of chromosome positioning is transmitted from mother to daughter cells during the second cleavage (Fig. 6F). While the mechanisms are not fully clear, this memory may reflect minimal relative repositioning of chromosomes during congression to and departure from the metaphase plate, perhaps due to the Rabl configuration and short cell cycles of the early embryo. This mitotic heritability of global chromosome positioning may influence processes that affect the viability and phenotype of the developing organism, such as rates of homologous recombination (HR)-mediated double-strand break (DSB) repair (*68*, *69*) and the distribution of translocations (*70*–*72*) in the early embryo.

Discussion

In this work we present in situ genome sequencing, unifying sequencing and imaging of genomes in intact samples. This unified approach enables de novo discovery of spatial organization of genomes across length scales, from single-cell sub-chromosomal domains to intercellular relationships. Because IGS is both sequencing and imaging-based, it can be extended in either modality based on the needs of specific experiments. We demonstrate this in early mouse embryos through integration of genotype information to spatially resolve the maternal and paternal genomes, through integration of immunofluorescence to localize genomic loci relative to nuclear landmarks, and by using whole-embryo spatial information to infer clonal lineages. This contextual information enabled us to uncover single-cell chromatin domains in zygotes with lamin-distal boundary positions and lamin-proximal interiors, as well as heritable correlations in global chromosome positioning within single early embryos.

While extant methods such as multiplexed DNA FISH and single-cell Hi-C are well-suited to measuring aspects of spatial genome organization, they cannot currently be combined in the same cell. With IGS, we spatially localize hundreds to thousands of genomic loci in single cells, achieving genomic resolutions comparable to recent genome-wide approaches based on targeted DNA FISH in fibroblasts (*73*, *74*). Unlike these targeted methods, we further show how IGS can perform untargeted sequence localization, resolve genome structure of maternal and paternal alleles, and be applied in 3D nuclei and thick intact samples. However, due to its genome-wide sampling frequency (at most ~1 Mb in this report), IGS is currently limited in its ability to systematically examine specific genetic loci in specific cells. Targeted DNA FISH or single-cell Hi-C are thus currently more appropriate for applications requiring high-resolution interrogation of genomic features such as TAD boundaries (*22*, *47*, *75*) or enhancer-promoter loops (*16*, *26*), and chromosome painting methods (*76*) may be preferred when high-throughput visualization of chromosomes territories is required. IGS therefore joins an ecosystem of conceptually new approaches (*77*–*79*)

complementary to these more well-established methods. We expect that improvements to IGS will further enable the study of genome sequence, structure, and function. In addition to the cultured cells and early embryos presented here, we anticipate extension of IGS to a broader range of cell types and intact tissues. Outside of development, IGS may be well-suited to study cancers, in particular how copy number instability and translocations contribute to tumor heterogeneity and alter nuclear morphology (*80*). The transposase-based library construction used in IGS may also be extended to measure the spatial localization of the accessible genome (*30*, *81*). Further, because nuclear volume is the primary constraint on the amplicon yield of IGS, we anticipate many-fold improvements in yield and resolution, either through smaller amplicons (*82*), or preferably through integration of IGS with Expansion Microscopy (*83*, *84*), which simultaneously increases nuclear volume by 50-100-fold and enables super-resolved imaging by physical expansion of samples. Finally, following our proof-of-concept integration with immunostaining, we expect increasingly multiplexed multiomic (*85*) variations of IGS will be possible. We anticipate in situ genome sequencing will be instrumental in unifying genomics and microscopy, and therefore sequence and structure.

Methods Summary
Library construction
Cells were grown and fixed on a glass coverslip, and embryos were fixed in solution and immobilized in polyacrylamide gel in a 6-well plate. Phosphorylated DNA adapters were inserted into fixed genomic DNA in situ by incubating samples 1 hour (cells) or overnight (embryos) with transposase. Hairpins were hybridized to the adapters on either side of the insert, and the complex was circularized by gap-fill ligation. Hairpins contained either a UMI and primers for in situ and ex situ sequencing, or an RCA primer hybridization site. RCA primers were hybridized, and RCA was performed overnight, with aminoallyl-dUTP spiked into the reaction. Amplicons were crosslinked by reacting with BSPEG(*9*).

In situ sequencing
For cells, the coverslip was mounted in a flow cell and in situ sequencing reactions were performed using automated fluidics. For embryos, reactions were performed manually in a 6-well plate. Samples were treated with calf intestinal phosphatase before the first primer hybridization and before each cleavage reaction. In situ sequencing was performed using sequencing-by-ligation chemistry. Samples were exchanged into an imaging buffer following each round of sequencing. Images were acquired using confocal microscopy. Immunostaining and immunofluorescence imaging were performed after in situ sequencing.

Ex situ sequencing
Samples were transferred into solution and used as input to a PCR reaction. The resulting library was sequenced using high-throughput paired-end sequencing and then then aligned to hg38 for PGP1f and mm10 for the early mouse embryos. Aligned reads that overlapped alleles that differ between the B6C3F1 and B6D2F1 strains were annotated as haplotype-informative.

Computational integration of *in situ and* ex situ *sequencing data*
In situ and ex situ sequencing data were computationally integrated using a probabilistic UMI matching approach that borrows principles from single-bit error correction to account for signal intermingling between densely-packed amplicons, as well as the decay of sequencing quality over successive rounds of in situ sequencing.

Detailed methods for sample preparation, library construction, multimodal sequencing, imaging, image analysis, and all analyses of data are described in the Materials and Methods section of the supplementary material. Oligonucleotide sequences are found in table S4. A description of methods for kit-free synthesis of in situ sequencing reagents is given in the supplementary text and tables S5 and S6. A description of cost, complexity and throughput of the method is found in the supplementary text and table S7.

REFERENCES AND NOTES
1. B. Bonev, G. Cavalli, Organization and function of the 3D genome. *Nat. Rev. Genet.* **17**, 661–678 (2016). doi:10.1038/nrg.2016.112 Medline
2. M. J. Rowley, V. G. Corces, Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* **19**, 789–800 (2018). doi:10.1038/s41576-018-0060-8 Medline
3. H. Zheng, W. Xie, The role of 3D genome organization in development and cell differentiation. *Nat. Rev. Mol. Cell Biol.* **20**, 535–550 (2019). doi:10.1038/s41580-019-0132-4 Medline
4. M. Spielmann, D. G. Lupiáñez, S. Mundlos, Structural variation in the 3D genome. *Nat. Rev. Genet.* **19**, 453–467 (2018). doi:10.1038/s41576-018-0007-0 Medline
5. J. Dekker, K. Rippe, M. Dekker, N. Kleckner, Capturing chromosome conformation. *Science* **295**, 1306–1311 (2002). doi:10.1126/science.1067799 Medline
6. E. Lieberman-Aiden, N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, J. Dekker, Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009). doi:10.1126/science.1181369 Medline
7. J. M. Levsky, R. H. Singer, Fluorescence in situ hybridization: Past, present and future. *J. Cell Sci.* **116**, 2833–2838 (2003). doi:10.1242/jcs.00633 Medline
8. J. R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, B. Ren, Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012). doi:10.1038/nature11082 Medline
9. T. Sexton, E. Yaffe, E. Kenigsberg, F. Bantignies, B. Leblanc, M. Hoichman, H. Parrinello, A. Tanay, G. Cavalli, Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**, 458–472 (2012). doi:10.1016/j.cell.2012.01.010 Medline
10. S. S. P. Rao, M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, E. L. Aiden, A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014). doi:10.1016/j.cell.2014.11.021 Medline

11. E. P. Nora, B. R. Lajoie, E. G. Schulz, L. Giorgetti, I. Okamoto, N. Servant, T. Piolot, N. L. van Berkum, J. Meisig, J. Sedat, J. Gribnau, E. Barillot, N. Blüthgen, J. Dekker, E. Heard, Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012). doi:10.1038/nature11049 Medline

12. R. Kalhor, H. Tjong, N. Jayathilaka, F. Alber, L. Chen, Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. *Nat. Biotechnol.* **30**, 90–98 (2011). doi:10.1038/nbt.2057 Medline

13. T. Nagano, Y. Lubling, T. J. Stevens, S. Schoenfelder, E. Yaffe, W. Dean, E. D. Laue, A. Tanay, P. Fraser, Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502**, 59–64 (2013). doi:10.1038/nature12593 Medline

14. T. Nagano, Y. Lubling, C. Várnai, C. Dudley, W. Leung, Y. Baran, N. Mendelson Cohen, S. Wingett, P. Fraser, A. Tanay, Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature* **547**, 61–67 (2017). doi:10.1038/nature23001 Medline

15. V. Ramani, X. Deng, R. Qiu, K. L. Gunderson, F. J. Steemers, C. M. Disteche, W. S. Noble, Z. Duan, J. Shendure, Massively multiplex single-cell Hi-C. *Nat. Methods* **14**, 263–266 (2017). doi:10.1038/nmeth.4155 Medline

16. T. J. Stevens, D. Lando, S. Basu, L. P. Atkinson, Y. Cao, S. F. Lee, M. Leeb, K. J. Wohlfahrt, W. Boucher, A. O'Shaughnessy-Kirwan, J. Cramard, A. J. Faure, M. Ralser, E. Blanco, L. Morey, M. Sansó, M. G. S. Palayret, B. Lehner, L. Di Croce, A. Wutz, B. Hendrich, D. Klenerman, E. D. Laue, 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* **544**, 59–64 (2017). doi:10.1038/nature21429 Medline

17. L. Tan, D. Xing, C.-H. Chang, H. Li, X. S. Xie, Three-dimensional genome structures of single diploid human cells. *Science* **361**, 924–928 (2018). doi:10.1126/science.aat5641 Medline

18. T. Cremer, M. Cremer, Chromosome territories. *Cold Spring Harb. Perspect. Biol.* **2**, a003889 (2010). doi:10.1101/cshperspect.a003889 Medline

19. J. Fraser, I. Williamson, W. A. Bickmore, J. Dostie, An Overview of Genome Organization and How We Got There: From FISH to Hi-C. *Microbiol. Mol. Biol. Rev.* **79**, 347–372 (2015). doi:10.1128/MMBR.00006-15 Medline

20. S. Wang, J.-H. Su, B. J. Beliveau, B. Bintu, J. R. Moffitt, C. T. Wu, X. Zhuang, Spatial organization of chromatin domains and compartments in single chromosomes. *Science* **353**, 598–602 (2016). doi:10.1126/science.aaf8084 Medline

21. D. I. Cattoni, A. M. Cardozo Gizzi, M. Georgieva, M. Di Stefano, A. Valeri, D. Chamousset, C. Houbron, S. Déjardin, J.-B. Fiche, I. González, J.-M. Chang, T. Sexton, M. A. Marti-Renom, F. Bantignies, G. Cavalli, M. Nollmann, Single-cell absolute contact probability detection reveals chromosomes are organized by multiple low-frequency yet specific interactions. *Nat. Commun.* **8**, 1753 (2017). doi:10.1038/s41467-017-01962-x Medline

22. B. Bintu, L. J. Mateo, J.-H. Su, N. A. Sinnott-Armstrong, M. Parker, S. Kinrot, K. Yamaya, A. N. Boettiger, X. Zhuang, Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science* **362**, eaau1783 (2018). doi:10.1126/science.aau1783 Medline

23. G. Nir, I. Farabella, C. Pérez Estrada, C. G. Ebeling, B. J. Beliveau, H. M. Sasaki, S. D. Lee, S. C. Nguyen, R. B. McCole, S. Chattoraj, J. Erceg, J. AlHaj Abed, N. M. C. Martins, H. Q. Nguyen, M. A. Hannan, S. Russell, N. C. Durand, S. S. P. Rao, J. Y. Kishi, P. Soler-Vila, M. Di Pierro, J. N. Onuchic, S. P. Callahan, J. M. Schreiner, J. A. Stuckey, P. Yin, E. L. Aiden, M. A. Marti-Renom, C. T. Wu, Walking along chromosomes with super-resolution imaging, contact maps, and integrative modeling. *PLOS Genet.* **14**, e1007872 (2018). doi:10.1371/journal.pgen.1007872 Medline

24. Q. Szabo, D. Jost, J.-M. Chang, D. I. Cattoni, G. L. Papadopoulos, B. Bonev, T. Sexton, J. Gurgo, C. Jacquier, M. Nollmann, F. Bantignies, G. Cavalli, TADs are 3D structural units of higher-order chromosome organization in *Drosophila. Sci. Adv.* **4**, eaar8082 (2018). doi:10.1126/sciadv.aar8082 Medline

25. E. H. Finn, G. Pegoraro, H. B. Brandão, A.-L. Valton, M. E. Oomen, J. Dekker, L. Mirny, T. Misteli; Extensive Heterogeneity and Intrinsic Variation in Spatial Genome Organization, Extensive heterogeneity and intrinsic variation in spatial genome organization. *Cell* **176**, 1502–1515.e10 (2019). doi:10.1016/j.cell.2019.01.020 Medline

26. L. J. Mateo, S. E. Murphy, A. Hafner, I. S. Cinquini, C. A. Walker, A. N. Boettiger, Visualizing DNA folding and RNA in embryos at single-cell resolution. *Nature* **568**, 49–54 (2019). doi:10.1038/s41586-019-1035-4 Medline

27. A. M. Cardozo Gizzi, D. I. Cattoni, J.-B. Fiche, S. M. Espinola, J. Gurgo, O. Messina, C. Houbron, Y. Ogiyama, G. L. Papadopoulos, G. Cavalli, M. Lagha, M. Nollmann, Microscopy-based chromosome conformation capture enables simultaneous visualization of genome organization and transcription in intact organisms. *Mol. Cell* **74**, 212–222.e5 (2019). doi:10.1016/j.molcel.2019.01.011 Medline

28. B. J. Beliveau, A. N. Boettiger, M. S. Avendaño, R. Jungmann, R. B. McCole, E. F. Joyce, C. Kim-Kiselak, F. Bantignies, C. Y. Fonseka, J. Erceg, M. A. Hannan, H. G. Hoang, D. Colognori, J. T. Lee, W. M. Shih, P. Yin, X. Zhuang, C. T. Wu, Single-molecule super-resolution imaging of chromosomes and in situ haplotype visualization using Oligopaint FISH probes. *Nat. Commun.* **6**, 7147 (2015). doi:10.1038/ncomms8147 Medline

29. Materials and methods are available as supplementary materials.

30. X. Chen, Y. Shen, W. Draper, J. D. Buenrostro, U. Litzenburger, S. W. Cho, A. T. Satpathy, A. C. Carter, R. P. Ghosh, A. East-Seletsky, J. A. Doudna, W. J. Greenleaf, J. T. Liphardt, H. Y. Chang, ATAC-see reveals the accessible genome by transposase-mediated imaging and sequencing. *Nat. Methods* **13**, 1013–1020 (2016). doi:10.1038/nmeth.4031 Medline

31. J. H. Lee, E. R. Daugharthy, J. Scheiman, R. Kalhor, J. L. Yang, T. C. Ferrante, R. Terry, S. S. F. Jeanty, C. Li, R. Amamoto, D. T. Peters, B. M. Turczyk, A. H. Marblestone, S. A. Inverso, A. Bernard, P. Mali, X. Rios, J. Aach, G. M. Church, Highly multiplexed subcellular RNA sequencing in situ. *Science* **343**, 1360–1363 (2014). doi:10.1126/science.1250212 Medline

32. J. R. Moffitt, J. Hao, G. Wang, K. H. Chen, H. P. Babcock, X. Zhuang, High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 11046–11051 (2016). doi:10.1073/pnas.1612826113 Medline

33. E. V. Volpi, E. Chevret, T. Jones, R. Vatcheva, J. Williamson, S. Beck, R. D. Campbell, M. Goldsworthy, S. H. Powis, J. Ragoussis, J. Trowsdale, D. Sheer, Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. *J. Cell Sci.* **113**, 1565–1576 (2000). Medline

34. A. Bolzer, G. Kreth, I. Solovei, D. Koehler, K. Saracoglu, C. Fauth, S. Müller, R. Eils, C. Cremer, M. R. Speicher, T. Cremer, Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLOS Biol.* **3**, e157 (2005). doi:10.1371/journal.pbio.0030157 Medline

35. A. P. J. de Koning, W. Gu, T. A. Castoe, M. A. Batzer, D. D. Pollock, Repetitive elements may comprise over two-thirds of the human genome. *PLOS Genet.* **7**, e1002384 (2011). doi:10.1371/journal.pgen.1002384 Medline

36. K. H. Miga, S. Koren, A. Rhie, M. R. Vollger, A. Gershman, A. Bzikadze, S. Brooks, E. Howe, D. Porubsky, G. A. Logsdon, V. A. Schneider, T. Potapova, J. Wood, W. Chow, J. Armstrong, J. Fredrickson, E. Pak, K. Tigyi, M. Kremitzki, C. Markovic, V. Maduro, A. Dutra, G. G. Bouffard, A. M. Chang, N. F. Hansen, A. B. Wilfert, F. Thibaud-Nissen, A. D. Schmitt, J.-M. Belton, S. Selvaraj, M. Y. Dennis, D. C. Soto, R. Sahasrabudhe, G. Kaya, J. Quick, N. J. Loman, N. Holmes, M. Loose, U. Surti, R. A. Risques, T. A. Graves Lindsay, R. Fulton, I. Hall, B. Paten, K. Howe, W. Timp, A. Young, J. C. Mullikin, P. A. Pevzner, J. L. Gerton, B. A. Sullivan, E. E. Eichler, A. M. Phillippy, Telomere-to-telomere assembly of a complete human X chromosome. *Nature* **585**, 79–84 (2020). doi:10.1038/s41586-020-2547-7 Medline

37. A. Cournac, R. Koszul, J. Mozziconacci, The 3D folding of metazoan genomes correlates with the association of similar repetitive elements. *Nucleic Acids Res.* **44**, 245–255 (2016). doi:10.1093/nar/gkv1292 Medline

38. V. Casa, D. Gabellini, A repetitive elements perspective in Polycomb epigenetics. *Front. Genet.* **3**, 199 (2012). doi:10.3389/fgene.2012.00199 Medline

39. M. Hausmann, J.-H. Lee, A. Sievers, M. Krufczik, G. Hildenbrand, COMBinatorial Oligonucleotide FISH (COMBO-FISH) with uniquely binding repetitive DNA probes. *Methods Mol. Biol.* **2175**, 65–77 (2020). doi:10.1007/978-1-0716-0763-3_6 Medline

40. W. Bao, K. K. Kojima, O. Kohany, Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob. DNA* **6**, 11 (2015). doi:10.1186/s13100-015-0041-9 Medline

41. M. Falk, Y. Feodorova, N. Naumova, M. Imakaev, B. R. Lajoie, H. Leonhardt, B. Joffe, J. Dekker, G. Fudenberg, I. Solovei, L. A. Mirny, Heterochromatin drives compartmentalization of inverted and conventional nuclei. *Nature* **570**, 395–399 (2019). doi:10.1038/s41586-019-1275-3 Medline

42. H. Muller, J. Gil Jr., I. A. Drinnenberg, The impact of centromeres on spatial genome architecture. *Trends Genet.* **35**, 565–578 (2019). doi:10.1016/j.tig.2019.05.003 Medline

43. C. B. Hug, J. M. Vaquerizas, The birth of the 3D genome during early embryonic development. *Trends Genet.* **34**, 903–914 (2018). doi:10.1016/j.tig.2018.09.002 Medline

44. C. Chazaud, Y. Yamanaka, Lineage specification in the mouse preimplantation embryo. *Development* **143**, 1063–1074 (2016). doi:10.1242/dev.128314 Medline

45. S. H. Namekawa, B. Payer, K. D. Huynh, R. Jaenisch, J. T. Lee, Two-step imprinted X inactivation: Repeat versus genic silencing in the mouse. *Mol. Cell. Biol.* **30**, 3187–3205 (2010). doi:10.1128/MCB.00227-10 Medline

46. M. Borsos, S. M. Perricone, T. Schauer, J. Pontabry, K. L. de Luca, S. S. de Vries, E. R. Ruiz-Morales, M.-E. Torres-Padilla, J. Kind, Genome-lamina interactions are established de novo in the early mouse embryo. *Nature* **569**, 729–733 (2019). doi:10.1038/s41586-019-1233-0 Medline

47. I. M. Flyamer, J. Gassler, M. Imakaev, H. B. Brandão, S. V. Ulianov, N. Abdennur, S. V. Razin, L. A. Mirny, K. Tachibana-Konwalski, Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature* **544**, 110–114 (2017). doi:10.1038/nature21711 Medline

48. Z. Du, H. Zheng, B. Huang, R. Ma, J. Wu, X. Zhang, J. He, Y. Xiang, Q. Wang, Y. Li, J. Ma, X. Zhang, K. Zhang, Y. Wang, M. Q. Zhang, J. Gao, J. R. Dixon, X. Wang, J. Zeng, W. Xie, Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* **547**, 232–235 (2017). doi:10.1038/nature23263 Medline

49. Y. Ke, Y. Xu, X. Chen, S. Feng, Z. Liu, Y. Sun, X. Yao, F. Li, W. Zhu, L. Gao, H. Chen, Z. Du, W. Xie, X. Xu, X. Huang, J. Liu, 3D chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis. *Cell* **170**, 367–381.e20 (2017). doi:10.1016/j.cell.2017.06.029 Medline

50. S. Collombet, N. Ranisavljevic, T. Nagano, C. Varnai, T. Shisode, W. Leung, T. Piolot, R. Galupa, M. Borensztein, N. Servant, P. Fraser, K. Ancelin, E. Heard, Parental-to-embryo switch of chromosome organization in early embryogenesis. *Nature* **580**, 142–146 (2020). doi:10.1038/s41586-020-2125-z Medline

51. M.-E. Torres-Padilla, D.-E. Parfitt, T. Kouzarides, M. Zernicka-Goetz, Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature* **445**, 214–218 (2007). doi:10.1038/nature05458 Medline

52. A. Burton, J. Muller, S. Tu, P. Padilla-Longoria, E. Guccione, M.-E. Torres-Padilla, Single-cell profiling of epigenetic modifiers identifies PRDM14 as an inducer of cell fate in the mammalian embryo. *Cell Rep.* **5**, 687–701 (2013). doi:10.1016/j.celrep.2013.09.044 Medline

53. A. Burton, M.-E. Torres-Padilla, Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. *Nat. Rev. Mol. Cell Biol.* **15**, 723–735 (2014). doi:10.1038/nrm3885 Medline

54. T. Aguirre-Lavin, P. Adenot, A. Bonnet-Garnier, G. Lehmann, R. Fleurot, C. Boulesteix, P. Debey, N. Beaujean, 3D-FISH analysis of embryonic nuclei in mouse highlights several abrupt changes of nuclear organization during preimplantation development. *BMC Dev. Biol.* **12**, 30 (2012). doi:10.1186/1471-213X-12-30 Medline

55. H. Fulka, F. Aoki, Nucleolus precursor bodies and ribosome biogenesis in early mammalian embryos: Old theories and new discoveries. *Biol. Reprod.* **94**, 143 (2016). doi:10.1095/biolreprod.115.136093 Medline

56. W. Mayer, A. Smith, R. Fundele, T. Haaf, Spatial separation of parental genomes in preimplantation mouse embryos. *J. Cell Biol.* **148**, 629–634 (2000). doi:10.1083/jcb.148.4.629 Medline

57. J. Reichmann, B. Nijmeijer, M. J. Hossain, M. Eguren, I. Schneider, A. Z. Politi, M. J. Roberti, L. Hufnagel, T. Hiiragi, J. Ellenberg, Dual-spindle formation in zygotes keeps parental genomes apart in early mammalian embryos. *Science* **361**, 189–193 (2018). doi:10.1126/science.aar7462 Medline

58. C. R. Cowan, P. M. Carlton, W. Z. Cande; Rabl Organization and the Bouquet, The polar arrangement of telomeres in interphase and meiosis. Rabl organization and the bouquet. *Plant Physiol.* **125**, 532–538 (2001). doi:10.1104/pp.125.2.532 Medline

59. M. Pouokam, B. Cruz, S. Burgess, M. R. Segal, M. Vazquez, J. Arsuaga, The Rabl configuration limits topological entanglement of chromosomes in budding yeast. *Sci. Rep.* **9**, 6795 (2019). doi:10.1038/s41598-019-42967-4 Medline

60. M. T. Lee, A. R. Bonneau, A. J. Giraldez, Zygotic genome activation during the maternal-to-zygotic transition. *Annu. Rev. Cell Dev. Biol.* **30**, 581–613 (2014). doi:10.1146/annurev-cellbio-100913-013027 Medline

61. K. Ahmed, H. Dehghani, P. Rugg-Gunn, E. Fussner, J. Rossant, D. P. Bazett-Jones, Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo. *PLOS ONE* **5**, e10531 (2010). doi:10.1371/journal.pone.0010531 Medline

62. N. Plachta, T. Bollenbach, S. Pease, S. E. Fraser, P. Pantazis, Oct4 kinetics predict cell lineage patterning in the early mammalian embryo. *Nat. Cell Biol.* **13**, 117–123 (2011). doi:10.1038/ncb2154 Medline

63. K. Piotrowska-Nitsche, A. Perea-Gomez, S. Haraguchi, M. Zernicka-Goetz, Four-cell stage mouse blastomeres have different developmental properties. *Development* **132**, 479–490 (2005). doi:10.1242/dev.01602 Medline

64. J. Walter, L. Schermelleh, M. Cremer, S. Tashiro, T. Cremer, Chromosome order in HeLa cells changes during mitosis and early G1, but is stably maintained during subsequent interphase stages. *J. Cell Biol.* **160**, 685–697 (2003). doi:10.1083/jcb.200211103 Medline

65. D. Gerlich, J. Beaudouin, B. Kalbfuss, N. Daigle, R. Eils, J. Ellenberg, Global chromosome positions are transmitted through mitosis in mammalian cells. *Cell* **112**, 751–764 (2003). doi:10.1016/S0092-8674(03)00189-2 Medline

66. J. Kind, L. Pagie, H. Ortabozkoyun, S. Boyle, S. S. de Vries, H. Janssen, M. Amendola, L. D. Nolen, W. A. Bickmore, B. van Steensel, Single-cell dynamics of genome-nuclear lamina interactions. *Cell* **153**, 178–192 (2013). doi:10.1016/j.cell.2013.02.028 Medline

67. I. Thomson, S. Gilchrist, W. A. Bickmore, J. R. Chubb, The radial positioning of chromatin is not inherited through mitosis but is established de novo in early G1. *Curr. Biol.* **14**, 166–172 (2004). doi:10.1016/j.cub.2003.12.024 Medline

68. C.-S. Lee, R. W. Wang, H.-H. Chang, D. Capurso, M. R. Segal, J. E. Haber, Chromosome position determines the success of double-strand break repair. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E146–E154 (2016). doi:10.1073/pnas.1523660113 Medline

69. E. R. Phillips, P. J. McKinnon, DNA double-strand break repair and development. *Oncogene* **26**, 7799–7808 (2007). doi:10.1038/sj.onc.1210877 Medline

70. E. Vanneste, T. Voet, C. Le Caignec, M. Ampe, P. Konings, C. Melotte, S. Debrock, M. Amyere, M. Vikkula, F. Schuit, J.-P. Fryns, G. Verbeke, T. D'Hooghe, Y. Moreau, J. R. Vermeesch, Chromosome instability is common in human cleavage-stage embryos. *Nat. Med.* **15**, 577–583 (2009). doi:10.1038/nm.1924 Medline

71. Y. Zhang, R. P. McCord, Y.-J. Ho, B. R. Lajoie, D. G. Hildebrand, A. C. Simon, M. S. Becker, F. W. Alt, J. Dekker, Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell* **148**, 908–921 (2012). doi:10.1016/j.cell.2012.02.002 Medline

72. J. M. Engreitz, V. Agarwala, L. A. Mirny, Three-dimensional genome architecture influences partner selection for chromosomal translocations in human disease. *PLOS ONE* **7**, e44196 (2012). doi:10.1371/journal.pone.0044196 Medline

73. H. Q. Nguyen, S. Chattoraj, D. Castillo, S. C. Nguyen, G. Nir, A. Lioutas, E. A. Hershberg, N. M. C. Martins, P. L. Reginato, M. Hannan, B. J. Beliveau, G. M. Church, E. R. Daugharthy, M. A. Marti-Renom, C. T. Wu, 3D mapping and accelerated super-resolution imaging of the human genome using in situ sequencing. *Nat. Methods* **17**, 822–832 (2020). doi:10.1038/s41592-020-0890-0 Medline

74. J.-H. Su, P. Zheng, S. S. Kinrot, B. Bintu, X. Zhuang, Genome-scale imaging of the 3D organization and transcriptional activity of chromatin. *Cell* **182**, 1641–1659.e26 (2020). doi:10.1016/j.cell.2020.07.032 Medline

75. J. M. Luppino, D. S. Park, S. C. Nguyen, Y. Lan, Z. Xu, R. Yunker, E. F. Joyce, Cohesin promotes stochastic domain intermingling to ensure proper regulation of boundary-proximal genes. *Nat. Genet.* **52**, 840–848 (2020). doi:10.1038/s41588-020-0647-9 Medline

76. B. D. Fields, S. C. Nguyen, G. Nir, S. Kennedy, A multiplexed DNA FISH strategy for assessing genome architecture in *Caenorhabditis elegans*. *eLife* **8**, e42823 (2019). doi:10.7554/eLife.42823 Medline

77. R. A. Beagrie, A. Scialdone, M. Schueler, D. C. A. Kraemer, M. Chotalia, S. Q. Xie, M. Barbieri, I. de Santiago, L.-M. Lavitas, M. R. Branco, J. Fraser, J. Dostie, L. Game, N. Dillon, P. A. W. Edwards, M. Nicodemi, A. Pombo, Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* **543**, 519–524 (2017). doi:10.1038/nature21411 Medline

78. S. A. Quinodoz, N. Ollikainen, B. Tabak, A. Palla, J. M. Schmidt, E. Detmar, M. M. Lai, A. A. Shishkin, P. Bhat, Y. Takei, V. Trinh, E. Aznauryan, P. Russell, C. Cheng, M. Jovanovic, A. Chow, L. Cai, P. McDonel, M. Garber, M. Guttman, Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* **174**, 744–757.e24 (2018). doi:10.1016/j.cell.2018.05.024 Medline

79. G. Girelli, J. Custodio, T. Kallas, F. Agostini, E. Wernersson, B. Spanjaard, A. Mota, S. Kolbeinsdottir, E. Gelali, N. Crosetto, M. Bienko, GPSeq reveals the radial organization of chromatin in the cell nucleus. *Nat. Biotechnol.* **38**, 1184–1193 (2020). Medline

80. D. Zink, A. H. Fischer, J. A. Nickerson, Nuclear structure in cancer cells. *Nat. Rev. Cancer* **4**, 677–687 (2004). doi:10.1038/nrc1430 Medline

81. J. D. Buenrostro, P. G. Giresi, L. C. Zaba, H. Y. Chang, W. J. Greenleaf, Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013). doi:10.1038/nmeth.2688 Medline

82. C.-M. Clausson, L. Arngården, O. Ishaq, A. Klaesson, M. Kühnemund, K. Grannas, B. Koos, X. Qian, P. Ranefall, T. Krzywkowski, H. Brismar, M. Nilsson, C. Wählby, O. Söderberg, Compaction of rolling circle amplification products increases signal integrity and signal-to-noise ratio. *Sci. Rep.* **5**, 12317 (2015). doi:10.1038/srep12317 Medline

83. F. Chen, P. W. Tillberg, E. S. Boyden, Optical imaging. Expansion microscopy. *Science* **347**, 543–548 (2015). doi:10.1126/science.1260088 Medline

84. S. Alon, D. R. Goodwin, A. Sinha, A. T. Wassie, F. Chen, E. R. Daugharthy, Y. Bando, A. Kajita, A. G. Xue, K. Marrett, R. Prior, Y. Cui, A. C. Payne, C.-C. Yao, H.-J. Suk, R. Wang, C.-C. Yu, P. Tillberg, P. Reginato, N. Pak, S. Liu, S. Punthambaker, E. P. R. Iyer, R. E. Kohman, J. A. Miller, E. S. Lein, A. Lako, N. Cullen, S. Rodig, K. Helvie, D. L. Abravanel, N. Wagle, B. E. Johnson, J. Klughammer, M. Slyper, J. Waldman, J. Jané-Valbuena, O. Rozenblatt-Rosen, A. Regev, IMAXT Consortium, G. M. Church, A. H. Marblestone, E. S. Boyden, Expansion sequencing: Spatially precise in situ transcriptomics in intact biological systems. bioRxiv 094268 [Preprint]. 15 May 2020. https://doi.org/10.1101/2020.05.13.094268.

85. C. Zhu, S. Preissl, B. Ren, Single-cell multimodal omics: The power of many. *Nat. Methods* **17**, 11–14 (2020). doi:10.1038/s41592-019-0691-5 Medline

86. E. Williams, J. Moore, S. W. Li, G. Rustici, A. Tarkowska, A. Chessel, S. Leo, B. Antal, R. K. Ferguson, U. Sarkans, A. Brazma, R. E. C. Salas, J. R. Swedlow, The Image Data Resource: A bioimage data integration and publication platform. *Nat. Methods* **14**, 775–781 (2017). doi:10.1038/nmeth.4326 Medline

87. A. C. Payne, Z. D. Chiang, P. L. Reginato, S. M. Mangiameli, zchiang/in_situ_genome_sequencing_processing: In situ genome sequencing processing code for: In situ genome sequencing resolves DNA sequence and structure in intact biological samples, Version 1.02, Zenodo (2020); https://doi.org/10.5281/zenodo.4299227.

88. A. C. Payne, Z. D. Chiang, P. L. Reginato, zchiang/in_situ_genome_sequencing_analysis: In situ genome sequencing analysis code for: In situ genome sequencing resolves DNA sequence and structure in intact biological samples, Version 1.0, Zenodo (2020); https://doi.org/10.5281/zenodo.4299125.

89. I. Solovei, A. Cavallo, L. Schermelleh, F. Jaunin, C. Scasselati, D. Cmarko, C. Cremer, S. Fakan, T. Cremer, Spatial preservation of nuclear chromatin architecture during three-dimensional fluorescence in situ hybridization (3D-FISH). *Exp. Cell Res.* **276**, 10–23 (2002). doi:10.1006/excr.2002.5513 Medline

90. C. A. Lareau, F. M. Duarte, J. G. Chew, V. K. Kartha, Z. D. Burkett, A. S. Kohlway, D. Pokholok, M. J. Aryee, F. J. Steemers, R. Lebofsky, J. D. Buenrostro, Droplet-based combinatorial indexing for massive-scale single-cell chromatin accessibility. *Nat. Biotechnol.* **37**, 916–924 (2019). doi:10.1038/s41587-019-0147-6 Medline

91. S. Picelli, Å. K. Björklund, B. Reinius, S. Sagasser, G. Winberg, R. Sandberg, Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res.* **24**, 2033–2040 (2014). doi:10.1101/gr.177881.114 Medline

92. J. H. Lee, E. R. Daugharthy, J. Scheiman, R. Kalhor, T. C. Ferrante, R. Terry, B. M. Turczyk, J. L. Yang, H. S. Lee, J. Aach, K. Zhang, G. M. Church, Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. *Nat. Protoc.* **10**, 442–458 (2015). doi:10.1038/nprot.2014.191 Medline

93. J. D. Buenrostro, B. Wu, H. Y. Chang, W. J. Greenleaf, ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* **109**, 21.29.1–21.29.9 (2015). doi:10.1002/0471142727.mb2129s109 Medline

94. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2013). doi:10.1038/nmeth.1923 Medline

95. T. Smith, A. Heger, I. Sudbery, UMI-tools: Modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* **27**, 491–499 (2017). doi:10.1101/gr.209601.116 Medline

96. T. Massingham, N. Goldman, Error-correcting properties of the SOLiD Exact Call Chemistry. *BMC Bioinformatics* **13**, 145 (2012). doi:10.1186/1471-2105-13-145 Medline

97. S. Berg, D. Kutra, T. Kroeger, C. N. Straehle, B. X. Kausler, C. Haubold, M. Schiegg, J. Ales, T. Beier, M. Rudy, K. Eren, J. I. Cervantes, B. Xu, F. Beuttenmueller, A. Wolny, C. Zhang, U. Koethe, F. A. Hamprecht, A. Kreshuk, ilastik: Interactive machine learning for (bio)image analysis. *Nat. Methods* **16**, 1226–1232 (2019). doi:10.1038/s41592-019-0582-9 Medline

98. H. R. Ueda, A. Ertürk, K. Chung, V. Gradinaru, A. Chédotal, P. Tomancak, P. J. Keller, Tissue clearing and its applications in neuroscience. *Nat. Rev. Neurosci.* **21**, 61–79 (2020). doi:10.1038/s41583-019-0250-1 Medline

99. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010). doi:10.1093/bioinformatics/btq033 Medline

100. N. Servant, N. Varoquaux, B. R. Lajoie, E. Viara, C.-J. Chen, J.-P. Vert, E. Heard, J. Dekker, E. Barillot, HiC-Pro: An optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **16**, 259 (2015). doi:10.1186/s13059-015-0831-x Medline

101. N. C. Durand, M. S. Shamim, I. Machol, S. S. P. Rao, M. H. Huntley, E. S. Lander, E. L. Aiden, Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **3**, 95–98 (2016). doi:10.1016/j.cels.2016.07.002 Medline

ACKNOWLEDGMENTS

(*88*) and a Github repository
(https://github.com/zchiang/in_situ_genome_sequencing_analysis/). Hi-C
data for PGP1f is available from the Gene Expression Omnibus (GEO) under
accession GSE12355. Hi-C data from early mouse embryos is available from GEO
accession GSE82185. DamID data from early mouse embryos is available from
GEO accession GSE112551. Genotypes for the B6C3F1 and B6D2F1 strains are
available from the Mouse Genomes Project
(https://www.sanger.ac.uk/data/mouse-genomes-project/). Mouse whole-
genome sequencing data are available from SRA accession SRR7511358.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/cgi/content/full/science.aay3446/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S27
Tables S1 to S7
References (*89–101*)
Movies S1 to S5
MDAR Reproducibility Checklist
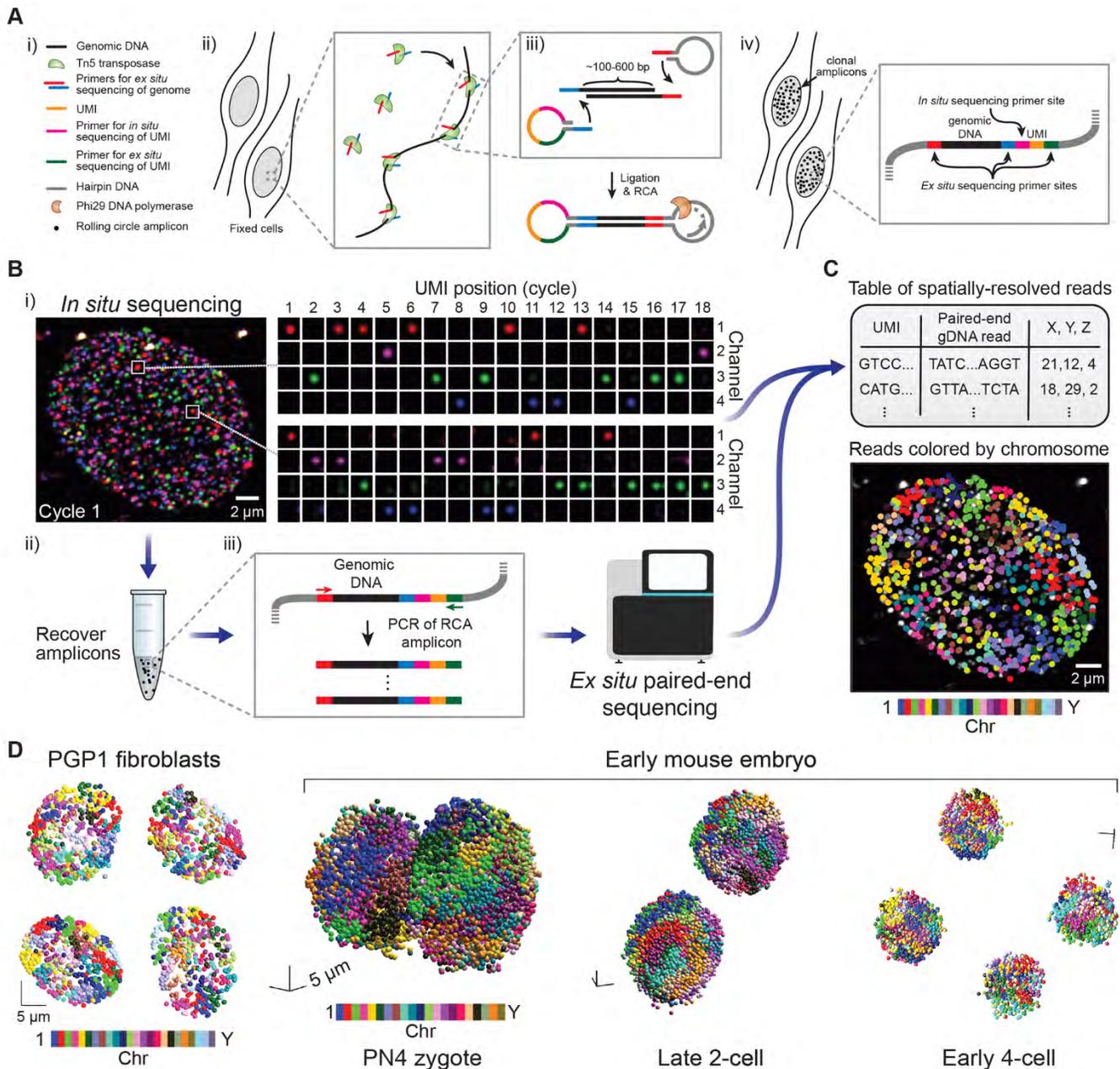
Fig. 1. Method for in situ genome sequencing. (A) In situ genomic DNA library construction. (i) Legend. (ii) Adaptor insertion. (iii) Insert circularization by hairpin ligation, followed by in situ rolling circle amplification (RCA). (iv) Clonal amplicons contain primers for in situ and ex situ sequencing. **(B)** Workflow for in situ genome sequencing. (i) In situ sequencing localizes unique molecular identifiers (UMIs). 4-channel imaging of two representative amplicons over 18 rounds of in situ sequencing. (ii) Amplicon dissociation following in situ sequencing. (iii) PCR and ex situ sequencing of amplicons associates genomic sequences with UMIs. **(C)** Top: paired-end sequences are spatially localized by integrating in situ and ex situ sequencing data. Bottom: matched reads, colored by chromosome, are overlaid on their imaged amplicon library (below). **(D)** In situ sequenced nuclei from cultured fibroblasts and intact embryos at the PN4 zygote, late 2-cell, and early 4-cell stages, with spatially localized reads colored by chromosome.
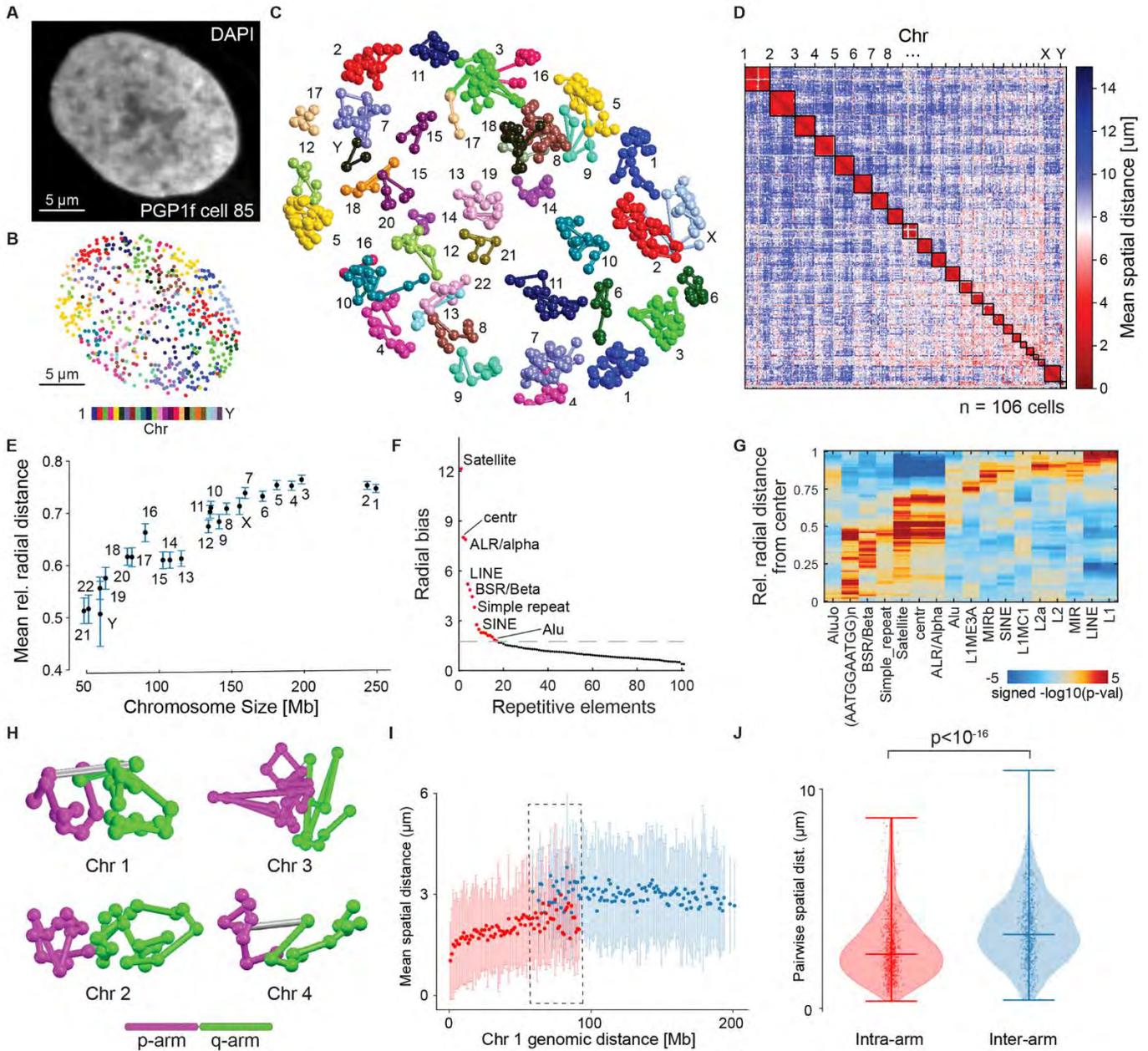
Fig. 2. IGS characterizes spatial features of the human genome. (A) DAPI stain of a PGP1f nucleus after in situ library construction. (B) 601 spatially localized reads in the same PGP1f nucleus, colored by chromosome. (C) Exploded view reveals conformations of chromosome territories, shown as in situ reads (balls) connected according to sequential genomic position (sticks). (D) Genome-wide population mean pairwise distance matrix of 106 PGP1f cells binned at 10 Mb. (E) Chromosome size vs. normalized mean radial distance from the nuclear center for 106 diploid-resolved PGP1f cells. Error bars denote 95% confidence interval of the mean determined by bootstrapping. (F) The 103 most abundant repetitive elements ordered by radial bias, defined as the variability of binned distances relative to a permuted background from the nuclear center for 106 PGP1f cells. The dashed grey line represents the threshold for elements shown in (G). (G) Radial enrichment/depletion by binned distance from the nuclear center for the repetitive elements with the strongest radial bias from (F). (H) Ball-and-stick models for Chr 1-4 in the same single-cell, demonstrating spatial polarization between the p and q arms of each chromosome. (I) Genomic distance vs. spatial distance for Chr 1, distinguishing intra-arm and inter-arm measurements. Error bars: standard deviation. Dashed: range in which both measurements can be compared at reasonable sampling depth (n > 20 per 1 Mb bin). (J) Intra-arm and inter-arm distance distributions in the dashed range in (I) are distributed differently (n = 819 intra-arm, 766 inter-arm, 144 Chr 1 territories, K-S test, p < 10^{-16}). Violin plot indicates median and range.
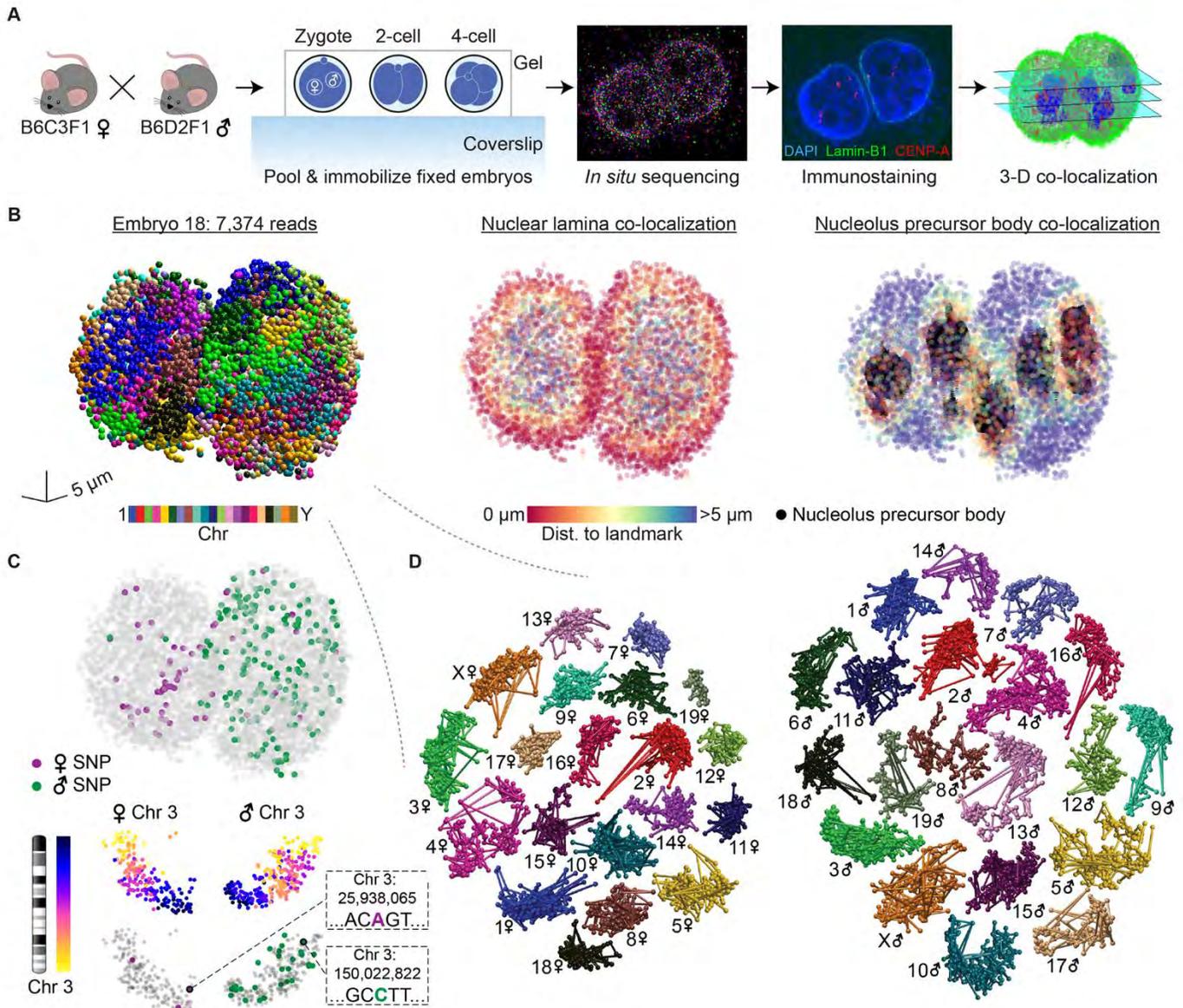
Fig. 3. IGS enables high-resolution genomic and spatial profiling of intact early mouse embryos. (A) Workflow. B6C3F1 x B6D2F1 embryos at the zygote, 2-cell, and 4-cell stages are pooled, fixed and immobilized in a polyacrylamide gel. Following in situ sequencing, DAPI and immunofluorescence staining of CENP-A and Lamin-B1 are performed. (B) Representative zygote with 7,374 spatially localized reads colored by chromosome (left), distance to the nuclear lamina (middle), and distance to nearest nucleolus precursor body (right). (C) Amplicons from (B), with reads colored by parental haplotype assignment for the intact embryo (top), reads colored by genomic position for Chr 3 homologs (middle), and reads colored by parental haplotype assignment for Chr 3 homologs (bottom). Boxes show two haplotype-informative Chr 3 SNPs. (D) An exploded view of chromosome territories from (B) for the maternal (left) and paternal pronuclei (right).

A Chromosome territories | Haplotype | Centromere-telomere position | GC content

Embryo 42, cell 1 of 2

1 ▮▮▮▮▮▮▮▮▮▮ Y   ♀ homolog   ♂ homolog   Cen ▮▮▮▮ Tel   35% ▮▮▮▮ 50%
Chr

Haplotype

B Embryo 42

Calculate haplotype separation →

♀ SNP   ♂ SNP

♀ ▮▮▮▮ ♂
1   0.5   1
Haplotype separation

C
Haplotype separation per cell

p<10⁻⁸   p<10⁻³

Zygote   2-cell   4-cell

More separate → Embryo 47, cell 3 (0.948)

Less separate → Embryo 54, cell 3 (0.824)

Centromere-telomere position

D Embryo 42, cell 1

90°

● CENP-A   Cen ▮▮▮▮ Tel

Chr 1

Chr 15

E
Mean neighbor centromere-telomere position

● Zygote
● 2-cell
● 4-cell

Centromere-telomere position

GC content

F Embryo 18

♀ Chr 12

♂ Chr 12

GC content   Dist. to lamina
35% ▮▮▮▮ 50%   0 µm ▮▮▮▮ >5 µm

G
Maternal zygote Chr 12 (ρ = 0.649)

Paternal zygote Chr 12 (ρ = 0.794)

Mean distance to lamina [µm]   GC content [%]

LADs

Genomic position on Chr 12 [1 Mb]

H
Per chromosome GC-lamina correlation [ρ]

p<10⁻⁵   n.s.

♀ ♂   ♀ ♂
Zygote   2-cell

Fig. 4. IGS characterizes developmental transitions in embryonic genome organization. (A) Exploded view of a single nucleus from a 2-cell embryo colored by chromosome territories, haplotype, centromere-telomere position, and GC content. (B) 2-cell embryo with spatially localized reads colored by parental haplotype assignment (left) and haplotype separation score (right). (C) Boxplots showing mean haplotype separation score per cell across developmental stages (left; K-S test, $p < 10^{-8}$ and $p < 10^{-3}$). Grey dots represent mean scores of single cells. Distribution mean (red line), 95% confidence interval (red box), and 1 standard deviation (blue box) are indicated. Two cells representing extreme scores (> 1 SD) are shown (right). (D) Nucleus from (A) with spatially localized reads colored by centromere-telomere position, shown from two angles 90 degrees apart (left). Black dots indicate the position of CENP-A as identified from immunostains. Chr 1 and Chr 15 homologs from this cell are shown (right) to illustrate the Rabl-like configuration. (E) Mean centromere-telomere position of spatial neighbors as a function of centromere-telomere position for each stage. (F) Chr 12 homologs from a representative zygote with spatially localized reads colored by GC content (left) and distance to lamina (right). (G) Plots showing the relationship between GC content and average distance to the nuclear lamina for 1 Mb bins in Chr 12 of the maternal and paternal zygotic pronuclei. Zygotic lamina-associated domains (LADs) defined by DamID are displayed below. (H) Boxplots showing Spearman's ρ between GC content and distance to lamina for 1 Mb bins, partitioned by haplotype and developmental stage (K-S test, $p < 10^{-5}$ and n.s.). Dots represent single chromosomes. Distribution mean (red line), 95% confidence interval (red box), and 1 standard deviation (blue box) are indicated. n = 24 zygotes, 40 2-cell, 49 4-cell nuclei for all panels. Scale bars: 5 μm in all directions.
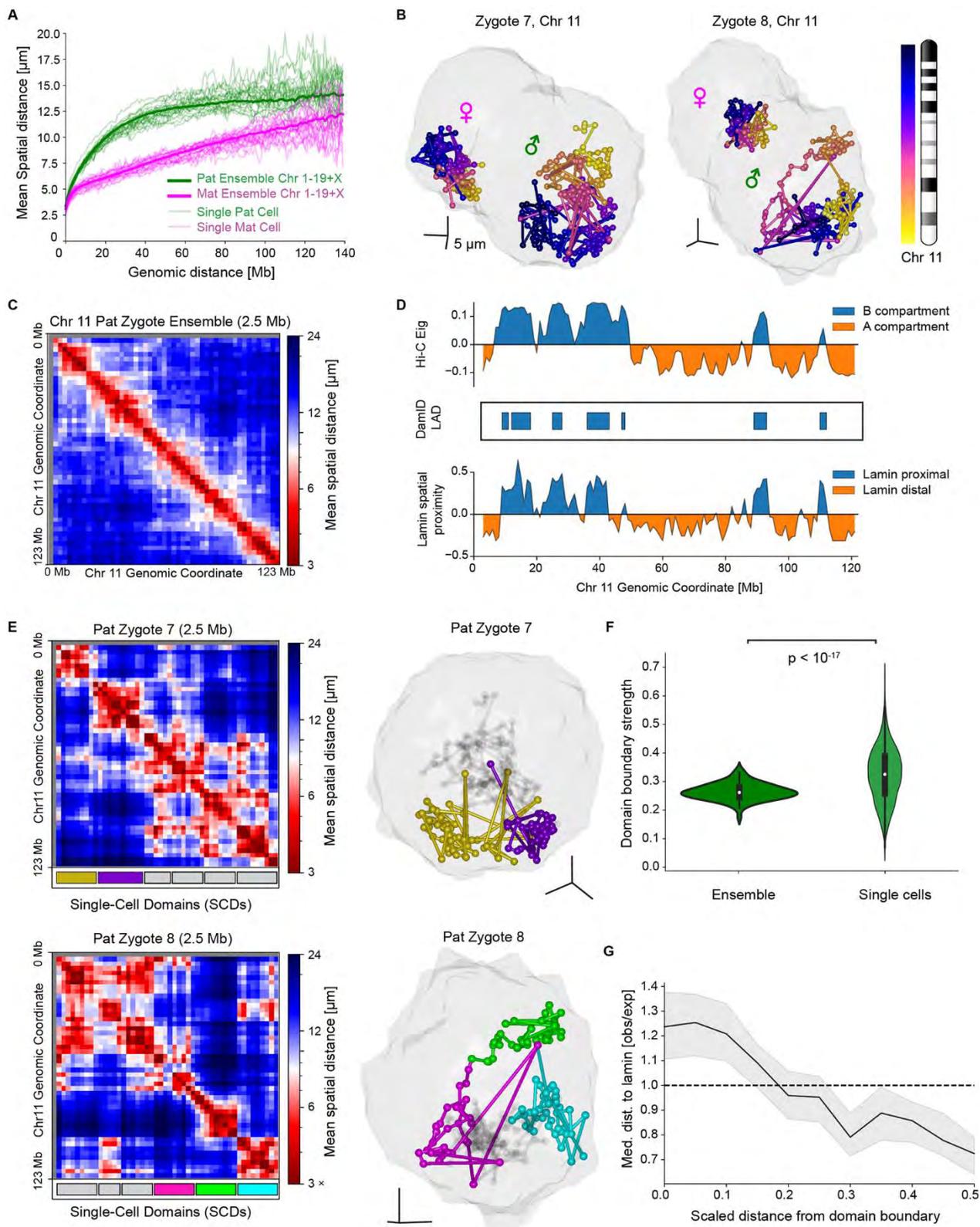
Fig. 5. IGS reveals single-cell domains in zygotes. (A) Global relationship between genomic and spatial distance in zygotes for all chromosomes, distinguishing the parental genomes. (B) Visualization of Chr 11 homologs in two zygotes according to parent-of-origin. (C) Population ensemble mean spatial distance matrix for paternal Chr 11, constructed at 2.5 Mb resolution (24 zygotic pronuclei, 2317 reads). (D) Comparison across measurement modalities for the population of paternal zygotic Chr 11. Top row: Hi-C defined eigenvalues and compartment calls. Middle row: DamID-defined population lamina associated domains. Bottom row: lamin-proximal and lamin-distal regions defined with IGS (24 zygotic pronuclei, 2317 reads). (E) Top left: single-cell mean distance matrix for paternal Chr 11 in a representative zygote, with single-cell domain boundaries (SCDs) marked below (263 reads). Top right: visualization of individual paternal SCDs in the same zygote. To assist visualization, two SCDs are shown in color (purple, gold), while the remaining SCDs are shown in grey. Bottom left: single-cell mean distance matrix for paternal Chr 11 in a second representative zygote, with single-cell domain boundaries marked below (213 reads). Bottom right: visualization of three paternal SCDs in the second zygote. To assist visualization, three SCDs are shown in color (magenta, lime, cyan) while the remaining SCDs are shown in grey. (F) Comparison of single-cell and ensemble domain boundary strengths spanning all detectable boundaries in Chr 1-19+X (74 ensemble boundaries, 1057 single-cell boundaries, K-S test, $p < 10^{-17}$). (G) Scaled distance from SCD boundary versus observed/expected median distance to nuclear lamina, measured genome-wide (Chr 1-19+X, N = 1262 SCDs). Envelope indicates 95% confidence interval determined by bootstrapping.
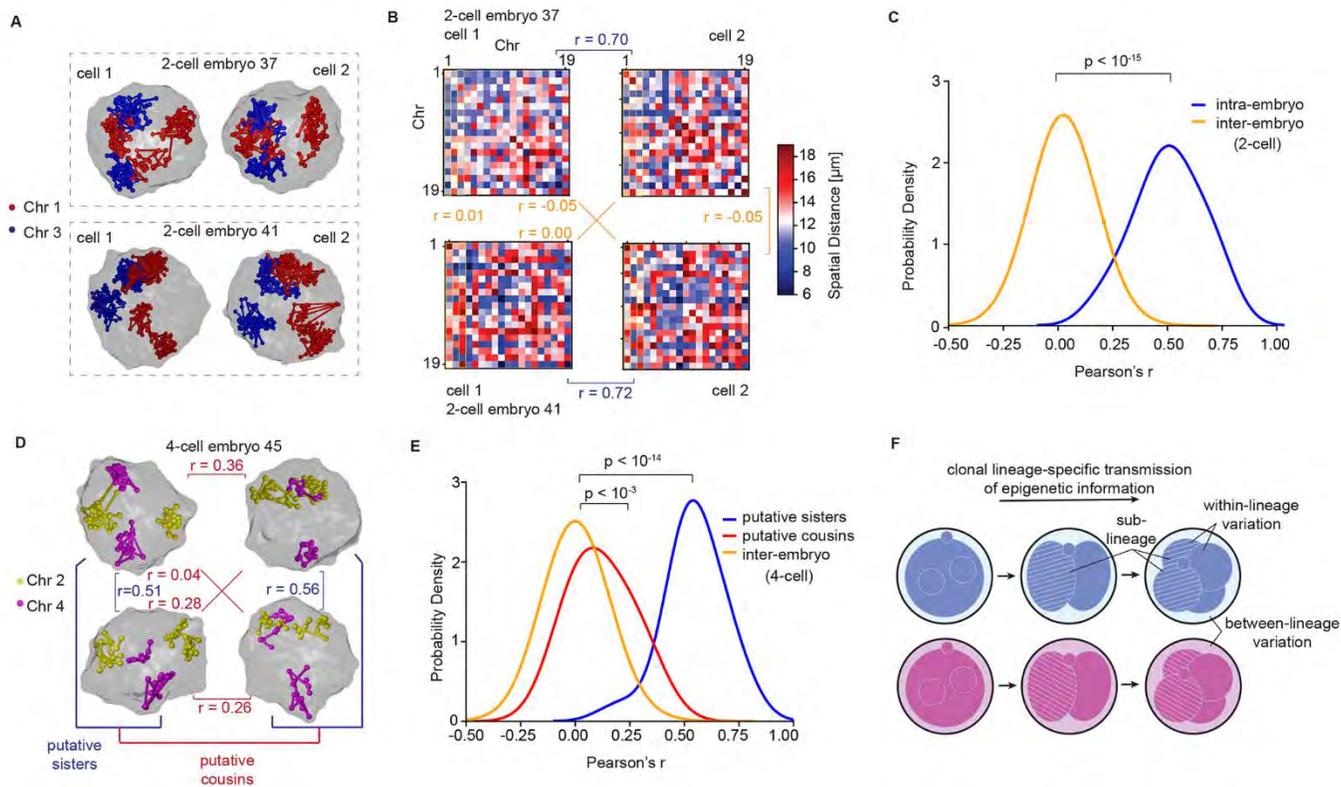
Fig. 6. IGS uncovers epigenetic memory of global chromosome positioning within single embryos. (**A**) Positioning of Chr 1 and 3 in the cells of 2-cell embryos 37 (top) and 41 (bottom). (**B**) Pairwise correlations between autosome distance matrices for the cells in (A). Intra-embryo and inter-embryo correlations are shown in blue and orange, respectively. (**C**) Probability distributions of correlations between autosome distance matrices for intra-embryo and inter-embryo pairs of cells among 2-cell embryos. K-S test, $p < 10^{-15}$; n = 20 intra-embryo pairs and n = 760 inter-embryo pairs, among 20 2-cell embryos. (**D**) Positioning of Chr 2 and 4 in the cells of 4-cell embryo 45. Pairs of cells are putatively classified as sister and cousin cells based on correlation of global chromosome positioning, with the most correlated pair classified as sisters. Correlations between sister and cousin cells are shown in blue and red, respectively. (**E**) Probability distributions of correlations between autosome distance matrices for pairs of putative sister cells, cousin cells, and inter-embryo pairs of cells among 4-cell embryos. K-S test, $p < 10^{-14}$ for sisters vs. inter-embryo and $p < 10^{-3}$ for cousins vs. inter-embryo; n = 18 sister pairs, n = 36 cousin pairs, and n = 933 inter-embryo pairs, among 13 4-cell embryos. (**F**) Model of epigenetic memory transmission within clonal lineages.

Science

In situ genome sequencing resolves DNA sequence and structure in intact biological samples

Andrew C. Payne, Zachary D. Chiang, Paul L. Reginato, Sarah M. Mangiameli, Evan M. Murray, Chun-Chen Yao, Styliani Markoulaki, Andrew S. Earl, Ajay S. Labade, Rudolf Jaenisch, George M. Church, Edward S. Boyden, Jason D. Buenrostro and Fei Chen

ARTICLE TOOLS	http://science.sciencemag.org/content/early/2020/12/29/science.aay3446
SUPPLEMENTARY MATERIALS	http://science.sciencemag.org/content/suppl/2020/12/29/science.aay3446.DC1
REFERENCES	This article cites 100 articles, 24 of which you can access for free http://science.sciencemag.org/content/early/2020/12/29/science.aay3446#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service

Supplementary Materials for

In situ genome sequencing resolves DNA sequence and structure in intact biological samples

Andrew C. Payne*, Zachary D. Chiang*, Paul L. Reginato*, Sarah M. Mangiameli, Evan M. Murray, Chun-Chen Yao, Styliani Markoulaki, Andrew S. Earl, Ajay S. Labade, Rudolf Jaenisch, George M. Church, Edward S. Boyden†‡, Jason D. Buenrostro†‡, Fei Chen†‡

*These authors contributed equally to this work.
†These authors contributed equally to this work.
‡Corresponding author. Email: edboyden@mit.edu (E.S.B.); jason_buenrostro@harvard.edu (J.D.B.); chenf@broadinstitute.org (F.C.)

This PDF file includes:

Materials and Methods
Supplementary Text
Figs. S1 to S27
Captions for Tables S1 to S4
Tables S5 to S7
Captions for Movies S1 to S5
References

Other Supplementary Material for this manuscript includes the following:
(available at science.sciencemag.org/cgi/content/full/science.aay3446/DC1)

Tables S1 to S4 (Excel)
Movies S1 to S5 (mp4)
MDAR Reproducibility Checklist (PDF)

Materials and Methods

Sample processing and library preparation

Experiments performed
Two separate IGS experiments were performed, with one experiment examining cultured fibroblasts and the other examining embryos.

Cell culture, fixation and permeabilization
PGP1f (Coriell GM23248) human primary skin cells were cultured to 70% confluence at 37°C with 5% CO_2 in 100 mm TC-Treated Culture Dishes (Corning) in Dulbecco's Modified Eagle Medium (DMEM), high glucose, GlutaMAX Supplement (Gibco 10566016), supplemented with 15% (v/v) fetal bovine serum (FBS) (Life Technologies, Inc 16140071), 1% (v/v) MEM Non-Essential Amino Acids Solution (Gibco 11140050), 100 units/mL penicillin and 100 µg/mL streptomycin. Cells were then treated with 0.25% Trypsin-EDTA (Gibco 25200056) for ~2 min at 37°C, mixed with culture media, centrifuged at 200xg for 5 min, and resuspended in 10 mL of fresh culture media. 125 µL of resuspended cells were then seeded into ethanol-sterilized, Matrigel-treated wells formed by 9 mm CultureWell silicone gaskets (Grace BioLabs 103240) attached to 40 mm circular coverslips (Bioptechs 40-1313-0319), except for the cells shown in **Fig. S2**, for which 90 µL of resuspended cells were seeded into Matrigel-treated sterile CultureWell chambered coverglass wells (Grace BioLabs 112358). Ethanol sterilization was performed by incubating the well with 70% ethanol for 15 min followed by aspiration and air-drying for 30 min. Matrigel treatment was performed by incubating the well with Matrigel Matrix, LDEV-free (Corning) diluted 1:50 in culture medium for 30 min. Cells were incubated overnight to allow them to attach to the coverslip. Cells were then washed once with 1x phosphate-buffered saline (PBS) (Gibco 10010023) and fixed with methanol-free paraformaldehyde (Electron Microscopy Sciences 15710) diluted to 4% in PBS for 10 min, washed with PBS, permeabilized with 0.5% Triton-X 100 (Sigma 93443) in PBS for 10 min, and then washed twice with PBS. Fixed, permeabilized cells were stored at 4°C in PBS up to three weeks.

Embryo collection
Cryopreserved viable embryos resulting from a cross of B6D2F1/Hsd males (Envigo) and superovulated B6C3F1/Hsd females (Envigo) were purchased from Embryotech Laboratories, Inc. Zygotes, 2-cell, and 4-cell embryos were collected at 22 hours, 43 hours, and 49 hours post human chorionic gonadotropin injection, respectively, by Embryotech Laboratories, Inc.

Embryo thawing, removal of zona pellucida, and fixation
Embryos were thawed as per instructions from Embryotech Laboratories, Inc.: straws containing embryos cryopreserved in freezing medium were thawed at room temperature for 2 minutes and then incubated in a 37°C water-bath for 1 minute, after which they were gently pushed out of the straw into a droplet of EmbryoMax Advanced KSOM Medium (Sigma MR-101). Embryos were then transferred to another droplet of the same medium

under oil and incubated for 1 hour at 37˚C with 5% CO_2 to recover from cryopreservation. The embryos were then briefly incubated in EmbryoMax Acidic Tyrode's solution (Sigma MR-004-D) to remove the zona pellucida and then transferred to a droplet of M2 medium (Sigma M7167). The zona-free embryos were then fixed in methanol-free paraformaldehyde 15710) diluted to 4% in PBS containing 1% polyvinylpyrrolidone, MW 360 kD (K90) (VWR AAJ61381-30) for 10 minutes at room temperature, followed by two rinses and final storage in PBS containing 1% polyvinylpyrrolidone, MW 360 kD (K90). Embryos were embedded in 4% polyacrylamide in a glass-bottom plate within 1 hr of fixation, to keep them immobilized during library preparation and *in situ* sequencing.

Embedding of embryos in polyacrylamide gel
The inner surface of a glass-bottom plate was incubated for 1 minute with 3-(Trimethoxysilyl)propyl methacrylate (Sigma 440159) diluted 1:1000 in 80% ethanol, 2% acetic acid, 18% H_2O, and then washed 3 times in ethanol to functionalize the glass surface with methacrylate moieties. A cut piece of a glass microscope slide, which would be used in casting the polyacrylamide gels, was incubated with Sigmacote (Sigma SL2) for 1 minute and then washed 3 times in H_2O to make the surface hydrophobic. Near each edge of the hydrophobic glass, two layers of Invisible Tape (Universal 83412) were attached to the glass to form spacers roughly 100 μm thick. Fixed embryos were quickly transferred through 3 droplets of polyacrylamide gel monomer solution (4% acrylamide (VWR 97064-870), 0.5% N,N′-diallyltartramide (Sigma 156868), 1X PBS, 0.1% Tween-20 (Sigma 11332465001), 0.1% ammonium persulfate (Sigma A3678), 0.1% N,N,N′,N′-tetramethylethylenediamine (Sigma T7024), 1:100 dilution of Fluorescent Tetraspeck 0.2 μm beads (Life Technologies, T7280)) and then transferred to the 6-well plate along with a small droplet of monomer solution taken from the final droplet. The hydrophobic glass slide with tape spacers was placed over the droplet to confine the monomer solution and embryos to a thin layer. The glass-bottom plate was placed in an air-tight container with a few ml of water in the bottom, which was then purged of oxygen for two minutes by placing the needle of a N_2 gas line into one of two holes poked in the lid of the container. The holes in the lid of the container were sealed, and the container was incubated at 37˚C for 2 hours. The hydrophobic glass was removed from the gel using tweezers, and the gel was washed in PBS.

Permeabilization of embryos
Following embedding in polyacrylamide gel, the embryos were incubated in 0.5% Triton-X 100 in PBS for 15 minutes, rinsed once and then washed 3 times for 5 minutes in PBS.

Tn5 purification
Tn5 was purified as previously described (*91*). E. coli cells (NEB C3013) harboring pTBX1-Tn5 were grown in terrific broth to an OD of 0.65 before addition of IPTG at 0.25 mM. Tn5 expression was induced at 23°C for 16-20 hours before harvesting by centrifugation and storage at -80°C until purification. 20 g of thawed E. coli pellet was lysed in 200 mL HEGX buffer (20 mM HEPES-KOH ph 7.2, 800 mM NaCl, 1 mM EDTA, 0.2% Triton, 10% glycerol) with cOmplete protease inhibitor (Roche) and 10 uL of benzonase (Thermo-Fisher Scientific). Cells were lysed using a LM20 microfluidizer device (Microfluidics) and cleared by centrifugation at 9000 x g for 30 min. 5.25 mL of

10% PEI (pH 7) was added dropwise to a stirring lysate solution to remove E. coli DNA and the resulting precipitation was removed by centrifugation for 10 min. Cleared supernatant was added to 30 mL of washed and equilibrated chitin resin (NEB), and mixed by end-over-end at 4°C for 30 min. Resin was washed by gravity flow with 1L HEGX buffer. To elute Tn5, 75 mL HEGX buffer with 100 mM DTT was added to the column, and 30 mL was drawn through the resin before sealing the column and storing at 4°C for 48h to allow for intein cleavage and elution of free Tn5. Eluted Tn5 was dialyzed into 2x Tn5 dialysis buffer (100 HEPES, 200 NaCl, 2 EDTA, 0.2 Triton, 20% glycerol), with two exchanges of 1L of buffer. The final solution was concentrated to 50 mg/mL as determined by A280 absorbance (A280 = 1 =0.616 mg/mL =11.56 mM) and flash frozen in liquid nitrogen before storage at -80°C.

Loading Tn5 transposase
Adaptor sequences (**Table S4**) were annealed by resuspending the strands at 50 uM each in 10 mM Tris-HCl, pH 8.0, 50 mM NaCl, followed by a thermal ramp from 65°C to 25°C over 1 hr. Annealed adaptors were mixed 1:1 with glycerol and stored at -20°C until Tn5 loading. Tn5 transposase was loaded by combining the two adaptors, Tn5 dilution buffer (50 mM Tris-HCl pH 7.5, 0.1 mM EDTA, 100 mM NaCl, 0.1% NP-40, 1 mM DTT, 50% glycerol) and 14.8 µM Tn5 transposase at a ratio of 1:1:1:1 and incubating for 30 min at RT. Loaded Tn5 was stored at -20°C until library preparation. We note that while unloaded Tn5 transposase was provided to us as a gift from a colleague who produced and purified it in-house, it is commercially available from Lucigen as Cat # TNP92110.

Hairpin snap-cooling
Hairpins (**Table S4**) were resuspended at 1 µM total hairpin concentration in 4X saline sodium citrate buffer (SSC), and then snap-cooled by heating at 95°C for 90 sec and then incubating on ice for 30 min.

Library preparation
Cells and gel-embedded embryos were treated with 0.1N HCl for 5 min at room temperature (RT) and washed 3x with 1X PBS. For adaptor transposition, cells were incubated for 1 hr (or overnight, for the cells shown in **Fig. S2)** and embryos were incubated overnight at 37°C with loaded Tn5 diluted 1:20 in 0.3X PBS, 10 mM Tris pH 8.5, 5 mM $MgCl_2$. We estimate that our in-house purified Tn5 has 2- to 4-fold the activity of Tn5 available through Lucigen as Cat # TNP92110; titration of Tn5 concentration or adjustment of incubation time for transposition may be needed to determine the optimal Tn5 concentration. Samples were then rinsed once and washed 3 times for 15 min at 37°C with 50 mM EDTA, 0.01% SDS in 1X PBS, and then washed twice with 1X PBS. Hairpins were hybridized to the adaptors by incubating samples with 250 nM annealed R2 hairpin, 62.5 nM each of annealed R1 hairpins A-D (cells) or 15.625 nM each of annealed R1 hairpins 1-16 (embryos) in 4X SSC for 1.5 hr (cells) or 2 hr (embryos) at 37°C. Samples were then washed 3 times for 5 min in PBS. Gap-fill and ligation was performed by incubating samples with 0.5 U/µL (cells) or 0.1 U/µL (embryos, cells shown in **Fig S2**) Ampligase Thermostable DNA Ligase (Lucigen A32750), 0.2 U/µL (cells) or 0.04 U/µL (embryos, shown in **Fig S2**) Phusion High-Fidelity DNA Polymerase (New England

Biolabs M0530), 50 µM dNTPs, 50 mM KCl, 20% formamide in 1x Ampligase Reaction Buffer (Lucigen) for 30 min at 37°C followed by 15 min at 45°C. (Partway through the study, we found that the lower concentrations of Ampligase and Phusion performed equally well to the higher concentrations.) Samples were then washed 3 times for 5 minutes in PBS. Rolling circle amplification (RCA) primers (**Table S4**) were hybridized by incubating the cells with 0.5 µM RCA primer in 2x SSC, 30% formamide for 2.5 hr (cells) or 3 hr (embryos) at 37°C, followed by washing once for 30 minutes (cells) or twice for 15 minutes (embryos) at 37°C with 2x SSC, 30% formamide. Samples were then washed 3 times with PBS. RCA was performed by incubating cells overnight with 1 U/µL Phi29 DNA Polymerase in 1X Phi29 DNA Polymerase Reaction Buffer (New England Biolabs M0269), and incubating embryos and the cells shown in **Fig S2** overnight with EquiPhi29 DNA Polymerase (Thermo Scientific) in 1X EquiPhi29 DNA Polymerase Reaction Buffer with 1 mM DTT. RCA reactions for all samples included 250 µM dNTPs and 50 µM aminoallyl dUTP. (Following the PGP1f experiment, we found that amplicon yields were higher when using EquiPhi29 compared to standard Phi29.) Samples were then rinsed and washed 3 times for 5 minutes with PBS. The samples were then treated for 1 hr at RT with 18 mg/mL BS(PEG)9 (Thermo Scientific 21582) in 10% DMSO, 1X PBS, to cross-link the amplicons. For visualization, an oligonucleotide (**Table S2**) modified with Cy3 was hybridized to the amplicons by incubating the cells with 100 nM oligo in 10% formamide, 4X SSC. Fluorescent Tetraspeck 0.2 µm beads were diluted 1:100 in PBS and applied to cells for 2 min before aspirating and washing with PBS.

<u>*In situ* sequencing and immunostaining</u>

Automated in situ sequencing in PGP1f cells
For cultured cells, we performed automated *in situ* sequencing by integrating automated fluidics, controlled by MATLAB, with a spinning disk confocal microscope (see *Imaging*), controlled by NIS-Elements AR software. The fluidics system was composed of a modular valve positioner with HVXM 8-5 valve (Hamilton 36766), PTFE laboratory tubing (Finemech S1810-12), an FCS2 flow cell (Bioptechs 060319-2), and a peristaltic pump (Rainin RP-1). A National Instruments Data Acquisition (NI-DAQ, USB-6008) card was used to connect the pump and the valve positioner to a computer. Sequencing reagents that required cooling (ligation mix, CIP solution, cleave 2, imaging buffer) were stored in a Mini Dry Bath (Fisher Scientific) set to 4°C.

To prepare the sample for sequencing, the silicone gasket was removed from the 40 mm round coverslip containing the cells and library, and the coverslip was placed inside the flow cell and connected to the fluidics. The visualization oligo was removed from the amplicons by continuously flowing strip solution (80% formamide in H$_2$O) over the sample at 500 µL/min for 10 min, and the sample was washed by continuously flowing instrument buffer (SOLiD Buffer F, 1:10 diluted) over the sample at 1000 µL/min for 5 min. All automated washing in the protocol going forward was performed in this manner.

In situ sequencing was performed using SOLiD chemistry as described previously (*92*), with modification. (While we purchased the SOLiD reagents commercially, we note that

they are no longer being sold as a kit. However, equivalent reagents can be synthesized by Integrated DNA Technologies and equivalent buffers can be prepared as simple formulations (*73*) (**Note S1**)). Four rounds of *in situ* sequencing were performed using each of five primers (**Table S4**), for a total of 20 rounds. All steps were performed using the automated fluidics and imaging setup described above, at RT. First, the sample was treated for 10 min with Quick Calf Intestinal Alkaline Phosphatase (Quick CIP) (New England Biolabs, M0508) diluted 1:20 in 1x CutSmart Buffer (New England Biolabs, B7204), and then 70% of the dead volume of CIP solution in the flow cell and fluidic lines was pumped back into the stock tube, to save reagents. Next, the sample was incubated for 10 min with a mixture of N-0 sequencing primers A-D (**Table S1)** at 625 nM each in 5x SASC (0.75 M sodium acetate, 75 mM tri-sodium citrate, pH 7.5). Four rounds of sequencing were subsequently performed. Each round of sequencing was performed as follows, except on the fourth round, where steps 3 and 4 were excluded: 1) A ligation reaction was performed by incubating with ligation mix (12 U/µL Rapid T4 DNA Ligase (Enzymatics, L6030-HC-L), 1:40 dilution of SOLiD sequencing oligos (SOLiD, 4475669), 1x T4 DNA Ligase Buffer (Enzymatics, B603)) for 10 min. After the ligation reaction, 70% of the dead volume of ligation mix in the flow cell and fluidic lines was pumped back into the stock tube, to save reagents. The sample was washed. 2) Imaging buffer (SOLiD Buffer A) (SOLiD, 4463024) was pumped into the flow cell, and then the sample was imaged (see *Imaging*) and washed. Photobleaching was performed after imaging by illuminating with the 594 nm laser line for 25 seconds. 3) A dephosphorylation reaction was performed by treating the sample with Quick CIP, as described above. The sample was washed. 4) A cleave reaction was performed by treating the sample first with SOLiD buffer C (SOLiD, 4458932) for 6 min and then with SOLiD buffer B (SOLiD, 4463021) for 6 min. The sample was washed with instrument buffer.

The sequencing primers were then stripped by continuously flowing strip solution (80% formamide in H_2O) over the sample at 500 µL/min for 10 min, and a mixture of N-1 primers A-D (recessed by one base at the 5' end, as compared to the N-0 primers A-D) were hybridized to the sample. Sequencing, stripping, and primer hybridization were repeated using N-2, N-3 and N-4 primers A-D. For the N-3 and N-4 primers, the first sequenced base was not imaged, because it is determined by the sequence of the primer binding site and provides no additional information on top of the first sequenced base of primer N-2, which distinguishes between primer binding sites A-D.

The sample was then stained with 100 nM Cy3-modified visualization oligo in 10% formamide, 4X SSC, for 45 minutes at RT, and with 1 µg/mL DAPI in 1X PBS for 3 minutes at RT, and then imaged for use in downstream image registration and segmentation.

For stratified sequencing in **Fig. S2**, only one sequencing primer was used at a time, and the sequencing protocol was performed by hand.

In situ *sequencing in embryos*
For gel-embedded embryos, sequencing reactions were performed manually at RT. Before sequencing, the embryos were stained with Cy3-modified visualization oligo and DAPI as

described above for cultured cells and imaged for use in downstream image registration and segmentation. The visualization oligo was stripped by rinsing and washing 2 times for 5 minutes in strip solution, and the embryos were rinsed and washed 4 times for 5 minutes in instrument buffer. The embryos were then treated with CIP solution for 45 min, rinsed and washed 5 times for 5 minutes in instrument buffer.

In situ sequencing was performed on the embryos as described for cultured cells with the following modifications: i) between each step, washes consisted of a rinse followed by 3 washes for 5 minutes each in instrument buffer, except following incubation with ligation mix, for which 3 washes for 10 minutes each were performed, and except for CIP solution, for which 5 washes for 5 minutes each were performed to thoroughly remove chloride ions that can cause precipitation with the silver ions in the SOLiD buffer C solution; ii) for the five primer hybridizations, a pool of primers 1-16 [N-0, N-1, N-2, N-3, or N-4] was used, at a concentration of 500 nM each and incubated for 45 minutes; iii) ligation mix used a 1:400 of dilution SOLiD sequencing oligos and included 1 mg/mL bovine serum albumin (Thermo Scientific 15561020); iv) the ligation reaction was incubated for 90 minutes; v) the incubation with SOLiD buffer C was incubated twice for 5 minutes followed by incubation with SOLiD buffer B twice for 5 minutes; vi) the first base of sequencing was collected for primer N-3, which provided the information to distinguish primer binding sites 1-16; vii) GLOX imaging buffer with Trolox was used (see below); viii) imaging was performed using different software and a different microscope (see *Imaging*, below); ix) primer stripping was performed by incubating 2 times for 5 minutes in strip solution.

In order to image thick samples (i.e early embryos), we used a different imaging buffer with improved antifade properties. The imaging buffer used for *in situ* sequencing of embryos was prepared as follows: i) a GLOX stock solution was prepared by dissolving 70 mg of glucose oxidase (Sigma G2133) in 1 ml of 200 mM Tris, pH 8.0, 50 mM NaCl, gently mixing with 250 µL of 10-60 mg/mL catalase solution (Sigma C100), centrifuging at maximum speed for 1 minute, taking the supernatant, and storing at 4˚C for up to 2 days; ii) immediately before imaging, GLOX stock solution was diluted 1:50 in 100 mM Tris pH 8.0, 25 mM NaCl, 2 mM Trolox (Sigma 238813), 10% glucose. The well of the glass bottom plate to be imaged was completely filled with GLOX imaging buffer and sealed with an adhesive polypropylene film (VWR 60941-070) to prevent contact with oxygenated air.

Immunostaining of embryos
Following *in situ* sequencing, embryos were immunostained. Lamin-B1 was stained with mouse Lamin B-1 Antibody (B10) (Santa Cruz; cat #sc-374015). CENP-A was stained with rabbit CENP-A (C51A7) mAb (Cell Signaling Technology cat #2048). anti-Lamin-B1 and anti-CENP-A were jointly diluted 1:1000 and 1:100 respectively in 3% BSA / 1x PBS. Embryos were stained overnight at 4C. After staining, samples were washed 3x in 1x PBS and detected with goat anti-mouse Alexa Fluor 488 (Thermo Fisher A32723) and goat anti-rabbit Alexa Fluor 647 (Thermo Fisher A32733). Secondaries were jointly diluted 1:200 in 1x PBS and the sample was stained for 1 hour. After staining, samples were washed 3x in 1x PBS and imaged.

Imaging

Imaging for cultured cells was performed using a Yokogawa CSU-W1 confocal spinning disk with Borealis modification coupled to a Nikon Ti-E inverted microscope with a Zyla 4.2 PLUS sCMOS camera, controlled by NIS-Elements AR software. The lasers, power, and emission filters used to image the fluorophores were: 100 mW solid state 405 nm smart diode laser at 50% power with 450/50 filter for DAPI, 150 mW solid state OPSL 488 nm laser at 70% power with 525/50 filter for FITC, 100 mW solid state OPSL 560 nm laser at 100% power with 582/15 filter for Cy3 and Alexa 546, 100 mW solid state DPSS 594 nm laser at 100% power with 624/40 filter for Texas Red, and 110 mW solid state OPSL 642 nm laser at 70% power with 685/40 filter for Cy5. A 1.40 NA 60x Plan Apochromat Lambda oil immersion objective lens (Nikon) with 0.3 μm step size and 200 ms exposure time were used for all images.

Imaging for embryos was performed using an Andor CR-DFLY-201-40 confocal spinning disk coupled to a Nikon Ti-E inverted microscope with a Zyla 4.2 PLUS sCMOS camera, controlled by Andor Fusion 2.0 software. The lasers and emission filters used to image the fluorophores were: 100 mW solid state 405 nm laser at with 450/50 filter for DAPI, 150 mW solid state 488 nm with 525/50 filter for FITC/AlexaFluor 488, 150 mW OBIS LS solid state OPSL 561 nm laser with 582/15 filter for Cy3, 100 mW OBIS LS 594 nm OPSS laser with 631/36 filter for Texas Red, 150 mW OBIS LX solid state 637 nm laser with 676/37 filter for Cy5/AlexaFluor 647. The dichroic mirror Andor CR-DFLY-DMQD-01 (405/488/561/640) was used for imaging all fluorophores except Texas Red, for which Andor CR-DFLY-DMQD-04 (405-445/514/594/730) was used. Images were collected using a 1.15 NA CFI Apo Long Working Distance Lambda S 40XC water immersion objective lens (Nikon) with 0.4 μm step size. For imaging the hybridization probe in the Cy3 channel, the power was 100% and exposure was 200 ms. For *in situ* sequencing, the power levels were 50% for FITC, 100% for Cy3, 100% for Texas Red, and 100% for Cy5. The exposure times were adjusted for each ligation number to account for reduction in brightness across successive ligations on the same primer. For FITC, Cy3, Texas Red, and Cy5, exposure times were: 300 ms, 300 ms, 100 ms, and 100 ms for the first ligation; 450 ms, 450 ms, 150 ms, and 150 ms for the second ligation; 675 ms, 675 ms, 225 ms, and 225 ms for the third ligation; 1000 ms, 1000 ms, 340 ms, and 340 ms for the fourth ligation. For DAPI imaging power was 50% and exposure time was 100 ms. For antibody imaging, power and exposure times were 100% and 150 ms for AlexaFluor 488 (lamin B1), 100% and 100 ms for AlexaFluor 647 (CENP-A).

Ex situ sequencing

Dissociation and PCR amplification of amplicons

Following *in situ* sequencing of cultured cells, the coverslip was removed from the flow cell. We carefully cut the coverslip down to the size of the imaged region (~250 cells) using a diamond scribe, and placed the remaining fragment in a PCR tube. The coverslip was cut to fit the sample in a PCR tube. A conventional Illumina sequencing library was prepared from the sample as follows: we first fully immersed the trimmed coverslip fragment by adding 100 μL PCR mix (50 μL Phusion U Hot Start PCR Master Mix (Thermo F533), 0.5

µL of 100 uM P5 primer, 0.5 µL of 100 µM v2_Ad2.41 primer, 49 µL ultrapure water), and performed a first round of amplification by incubation at 8˚C for 10 min, 98˚C for 30 s, 10 cycles of [98˚C for 10 s, 69˚C for 15 s, 72˚C for 30 s], 72˚C for 5 min, 4˚C hold. To avoid overamplification, we then eluted 5 µL for quantification by qPCR, as described in (*93*), maintaining the rest of the sample at 4 C. We then performed an additional 15 cycles of amplification on the remaining 95 µL of sample as described above. Following PCR, the reaction was eluted from the PCR tube, column purified (DNA Clean and Concentrator-5, Zymo) using 475 µL DNA binding buffer (i.e. 1:5) and eluted into 20 µL water.

Following *in situ* sequencing of polyacrylamide-embedded embryos, the gels were washed with ultra pure water, scored with a razor, scraped off of the glass-bottom plate in pieces using a disposable polypropylene spatula, and transferred using a paintbrush (Blick Art Supplies, 06170-7030) to PCR tubes containing 34 ul of PCR mix each. PCR mix was prepared as follows: 20 µL Phusion U Hot Start PCR Master Mix, 1 µL 20 µM P5 primer, 2 µL 10 µM i7 indexed Nextera primer (Illumina FC-131-1002), 11 µL ultra pure H_2O. The gel fragments were assumed to contain 6 µL of ultra pure water. The tubes were flicked several times, then incubated on ice 3 times for 10 minutes with flicking between each incubation. The tubes were then pre-amplified by thermocycling according to the following program: 98˚C for 30 s, 5 cycles of [98˚C for 45 s, 72˚C for 80 s], 4˚C hold. The tubes were then flicked and incubated at 4˚C overnight to allow amplicons to diffuse out of the gel. The tubes were flicked again. 1.5 µL of the pre-amplified libraries was diluted 1:10 in PCR mix with 1x SYBR Green followed by quantification by qPCR. Half the remaining volume of the pre-amplified libraries was then diluted 1:10 in PCR mix and thermocycled according to the following program: 98˚C for 30 s, [# qPCR cycles required to reach ½-maximum + 1] cycles of [98˚C for 10 s, 72˚C for 40 s], 72˚C for 5 min, 4˚C hold. The number of cycles used in the latter PCR reaction ranged from 18-24. Dilution of the pre-amplified libraries was to dilute putative carry-over from the in situ sequencing, which dramatically improved yield.

Ex situ *Illumina sequencing*
PCR-amplified libraries from cultured cells were sequenced using one lane of a HiSeq 2500. 84 bp paired-end reads were collected with a 25 bp i5 index read. Data was analyzed using the Broad Picard Pipeline, which includes de-multiplexing and data aggregation. PCR-amplified libraries from embryos were sequenced on NovaSeq 6000. 143 bp paired-end reads were collected with a 21 bp i5 index read. Sequencing data was demultiplexed using bcl2fastq2 (v2.19.1).

Ex situ *sequence alignment and processing*
Ex situ sequenced UMIs were appended to the headers of associated paired-end sequencing reads. Reads were then trimmed for sequencing adapters using a custom python script and aligned to hg38 and mm10 using bowtie2 (*94*) with --very-sensitive and -k 5 parameters. The resulting BAM files were sorted by genomic coordinates, UMIs were moved from the header to a new read group tag, and optical duplicates were removed using Picard MarkDuplicates (http://broadinstitute.github.io/picard/). 76,879,813 reads were sequenced from PGP1f cells with a 96.11% alignment rate to hg38 and a 90.5% duplication rate.

986,237,536 reads were sequenced from embryos with a 97.94% alignment rate to mm10 and a 99.7% duplication rate.

For the PGP1f data, the filtered BAM file was split into three files -- one for uniquely-aligning reads, a second for multi-mapped reads, and a third for unmapped reads. UMIs for the uniquely-aligned and multi-mapped reads were then used to group PCR duplicates together using UMI-tools group with parameter --edit-distance-threshold 2 to facilitate UMI error correction (*95*). For groups in which the inferred true UMI was ambiguous, the highest quality base was chosen based on quality scores from original UMI FASTQ file instead of randomly, as implemented in UMI-tools. Multi-mapped and unmapped reads were collapsed into a single entry per UMI to be added to a comprehensive list of observed UMIs. This list was then filtered for occurrences of index swapping based on the frequency of PCR duplicates for each unique UMI-genomic location combination. Lastly, the UMIs were converted to colorspace sequences using a di-base encoding table (*96*).

Image processing and UMI matching

Field of view image processing

All image processing and UMI matching steps were implemented in custom MATLAB scripts available from the GitHub repository. For all fields of view, 3D image stacks for each sequencing cycle and DAPI stains were registered to cycle 1 using normalized cross-correlation to correct for shifts that may have occurred between imaging rounds. For the PGP1f data, the 3D image stacks for each channel were also corrected for spectral and physical drift by calculating a 3D rigid transformation based on the positions of fluorescent Tetraspeck beads, which were detected using a 3D peak finder. Bounds for each nucleus were defined by performing threshold-based segmentation on the DAPI image stack, and nuclei located at the edge of the field of view that were not fully imaged in every sequencing cycle were excluded from further analysis. In some (<10%) embryos, bounds were manually added for nuclei that were not automatically detected.

Nucleus image processing

A five-dimensional (*x* by *y* by *z* by channel by cycle) image stack was created for each nucleus by cropping fields of view based on nuclei segmentation bounds. Next, stacks were deconvolved using a high pass Gaussian filter to improve the resolution of densely-packed amplicons in the nucleus. The images in each stack were registered to each other along the cycle dimension using an iterative approach. For embryos, each channel from sequencing cycle 1 was independently registered to the visualization oligo image via a 3D affine transformation. For all data, images from a single cycle were collapsed across the channel dimension using a maximum intensity projection, and registered to the collapsed cycle 1. The resulting transformation was then applied to the uncollapsed images from all four channels separately, which were then re-registered independently to the collapsed cycle 1. This process was repeated for all cycles. In the PGP1f data, the rate of signal phasing between cycles was calculated by identifying pixels that lose most, but not all of their signal in subsequent cycles. The phasing rate was then calculated from these pixels and subtracted stack-wide to correct any residual fluorescence resulting from sequencing inefficiencies.

Lastly for all data, images from each cycle were normalized by applying quantile normalization such that the total fluorescence values from each channel were equal (**Fig. S4**).

Amplicon identification and size quantification
Amplicon centers were identified by applying a 3D peak finder to the normalized image stacks, with peaks under a percentile-based threshold being removed from downstream analysis. Peaks from separate images with identical 3D coordinates were collapsed into a single entry. To quantify the distribution of amplicon sizes in PGP1f, a Gaussian fit was simultaneously performed on every putative amplicon within a single nucleus (**Fig. S3**). The majority of amplicons had diameters between 400-500 nm, and in almost all cases, the identified peak was confirmed to be at the center of the amplicon. Each amplicon was thus localized to a single 0.108 by 0.108 by 0.3 or 0.4 μm voxel (0.3 in PGP1, 0.4 in embryos).

In situ *UMI processing*
For each amplicon, a region of interest was defined by selecting nearby pixels with high correlation to the peak over all images. Each region was then quantified over all channels and cycles by summing the fluorescence values of all pixels. The resulting two-dimensional matrix (channels by cycles) was normalized such that the sum of squares for each column = 1. Since a large fraction of amplicons fall in densely-packed subcellular volumes, regions were iteratively refined to maximize the purity score, which is calculated for a set of pixels by taking the negative log transform of the maximum fluorescence values across channels and multiplying the value from each cycle together i.e. -log10(product(max(matrix,1),2)). Following this iterative refinement, each amplicon was associated with a final colorspace probability matrix representing the region with the highest signal purity (**Fig. S5**).

Generating spatially-resolved reads via UMI matching
UMI matching was facilitated by a set of colorspace probability matrices derived from the spatially-resolved amplicons in the *in situ* images, and a list of observed UMIs and associated genomic reads from the *ex situ* sequencing. A consensus colorspace sequence was generated from each probability matrix by taking the channel with the maximum probability from each cycle. Each consensus sequence was then compared against all observed sequences using Hamming distance. All sequences with Hamming distances less than 4 were saved as potential matches. If there were no observed sequences with Hamming distance less than 4, the threshold was incremented by 1 until one or more sequences were found. For each potential matching sequence, a match score was calculated by tracing the path of the observed sequence though the colorspace probability matrix. Each value along the path was multiplied together and the final product was negative log transformed. Lower match scores indicate a more probable match, due to the negative log transformation. The potential matching sequence with the lowest match score was associated with each amplicon, generating spatially-resolved reads (**Fig. S5**). For the embryos, each nucleus was only allowed to contain matches from a single sequencing well. A consensus well was called for each nucleus, and any matched reads from a different well were excluded.

UMI match filtering

Spatially-resolved reads were filtered based on both purity score, a measure of how clearly an amplicon could be resolved in the images, and match score, a measure of how well the observed *ex situ* UMIs match the colorspace probability matrix. Reads with lower purity scores were allowed to have higher match scores, analogous to a continuous Hamming distance threshold. In order to calculate the rate of spatially-resolved amplicons that successfully match an *ex situ* genomic read, a set of high quality reads was defined by taking all reads with a purity score less than 1.5 in PGP1f and 0.75 in embryos. Matching rate was then calculated by dividing the number of high quality reads passing filter over the number of toal high quality reads (**Fig. S6**).

UMI matching false discovery rate

False discovery rate (FDR) was calculated independently for each putative amplicon. For each corresponding colorspace probability matrix, a $4^{18/19}$ tree representing all possible 18/19-base UMIs was generated (18 in PGP1f, 19 in embryos), with branches over the match threshold being pruned to save computational resources. The total number of possible UMIs that pass the match filter for each amplicon was then divided by $4^{18/19}$ and multiplied by the number of observed *ex situ* UMIs per sequencing well to yield a per-amplicon FDR.

Registration of immunofluorescence images

For the embryo data, immunofluorescence (IF) images of lamin B1 and CENP-A were collected along with an additional DAPI stain after sequencing. To register these images to the *in situ* sequencing images, this additional DAPI stain was registered to the original DAPI stain imaged at the time of sequencing to calculate a 3D affine transformation for each nucleus. This transformation was then applied to the IF images.

Segmentation of nuclear landmarks

For the embryo data, segmentation of nuclear landmarks was performed in one of two ways. The spatial location of the nuclear lamina was segmented from the lamin B1 immunofluorescence images using a percentile-based threshold. Both lamin B1 in the nuclear interior and nuclear lamina invaginations were included in the final segmentation. The spatial positions of nucleolus precursor bodies and centromeres were segmented from the DAPI stain and CENP-A immunofluorescence images respectively by training an object classification workflow for each developmental stage in Ilastik (*97*). Distance to these landmarks for each spatially-localized read was calculated by generating pairwise distances to all segmented voxels and finding the minimum spatial distance.

Data filtering and quality control

Cell filtering

For PGP1f, 88 cells with fewer than 200 spatially-localized reads and 27 cells with detected aneuploidy were excluded. This resulted in a dataset containing 106 cells with 328 ± 114 reads per cell (mean ± SD; **Table S1**), equivalent to an average of one genomic locus per ~18 Mb across each diploid human genome.

For embryos, 4 embryos with fewer than 600 spatially-localized reads were excluded. Embryos were visually inspected after DAPI staining, and 1 mitotic zygote and 1 PN5 zygote were identified and excluded. In one 2-cell embryo that was excluded for having fewer than 600 localized reads, both cells were also found to be mitotic. 1 haploid 2-cell embryo was identified based on visual inspection of data and excluded. Polar bodies were detected by visual inspection and removed. This resulted in a dataset containing 24 zygotes, 40 nuclei from 2-cell embryos, and 49 nuclei from 4-cell embryos, with $3,909 \pm 2,116$, $2,357 \pm 1,063$, and $1,074 \pm 622$ reads per nucleus (mean $\pm$ SD) for zygotes, 2-cell, and 4-cell embryos, equivalent to an average of one locus per 1.3 Mb, 2.1 Mb, and 4.7 Mb across each diploid mouse genome, respectively (**Table S2**). We note that for 3 4-cell embryos, *in situ* sequencing data could not be collected from one of the four cells due to z-dependent optical inhomogeneities. We expect that this effect will be mitigated in future experiments using optical clearing techniques (*98*).

Amplicon density inside and outside of nuclei
In cultured cells, amplicons were detected using peak detection in ImageJ for maximum projection images of *in situ* sequencing libraries stained with a visualization probe. A nuclear mask was produced in ImageJ using DAPI staining of nuclei, with a manually adjusted brightness threshold. Amplicons coinciding with the DAPI mask, or within 8 pixels of the mask, were counted as being associated with nuclei. The 8-pixel padding was used because amplicons at the periphery of the nucleus sometimes fell <1 μm outside the boundary of the DAPI mask. This analysis was performed for every second field of view in the dataset. Mean $\pm$ standard deviation amplicon density was $0.94 \pm 0.06/\mu m^2$ inside nuclei, and $0.006 \pm 0.001/\mu m^2$ outside nuclei.

Estimation of nuclear volume and density of spatially-localized reads
The volume of each nucleus was estimated by taking the number of voxels segmented in the DAPI stain and multiplying by the volume of a single voxel, 0.108 μm * 0.108 μm * 0.3 or 0.4 μm (0.3 in PGP1, 0.4 in early mouse embryos). The read density of each nucleus was calculated by dividing the number of spatially-localized reads by the estimated volume of the nucleus.

Comparison of genomic coverage to whole-genome sequencing
Spatially-localized reads were partitioned into 2.5 Mb (for PGP1f) and 1 Mb (for embryos) genomic bins. The number of reads in each bin was considered to be the raw coverage of each bin. Bins were normalized by calculating the GC content of each bin, and then dividing each by the mean of the 500 bins with the most similar GC content. The resulting values were then divided by the median normalized coverage of autosomal bins and multiplied by 2 to scale the coverage values to copy number. PGP1f whole-genome sequencing data for comparison was downloaded from ENCODE accession ENCFF713HUF) and subsampled to match the size of the PGP1f IGS data set. Whole-genome sequencing for mouse was obtained from SRR7511358 and subsampled to match the size of the embryo IGS data set. The data was then normalized for GC content and scaled to copy number in the same fashion as the IGS data.

Quantification of chromatin accessibility bias
A list of transcription start sites (TSSs) in mm10 was obtained from the UCSC Genome Browser (https://genome.ucsc.edu/cgi-bin/hgTables). All spatially-localized reads within 2000 bp of each TSS were identified. The starting positions of each identified read relative to the TSS were calculated and aggregated into bins. Enrichment was calculated by dividing the number of reads in each bin by the average of reads 1800-2000 bp upstream of the mean TSS. The ATAC-seq mouse brain data set for comparison to IGS was obtained from a previous study (*90*). The data was subsampled to match the size of the embryo IGS data set, and TSS enrichment was calculated in the same fashion.

Data annotation

Spatial clustering of reads in cultured cells
For cultured cells, amplicons were clustered and resolved into homologous chromosome pairs using an approach based on maximum likelihood estimation. Prior to clustering, we first identify and outlier points using DBSCAN, where the search radius parameter is set to 3.5 μm and a minimum of two neighbors is required to identify a core point. Cluster assignments were initially generated for all paired chromosomes using k-medoids (k=2). In many cases this was adequate to resolve the homologs, as they localized to different regions of the nucleus and could be clustered based on spatial coordinates alone; these cases were identified manually by visual inspection and used to infer the homolog assignments for the entire data set as follows. Using the subset of well-separated homologous chromosomes, we constructed the joint probability distribution of genomic and physical pairwise distances. The distribution was smoothed using a gaussian kernel and normalized over the length of the chromosome and the range of observed physical distances. To perform maximum likelihood estimation, we first initialized the cluster assignments using k-medoids. Starting with the amplicon having the smallest genomic position, we assign each segment along the chromosome (in the preliminary cluster) a probability based on the empirically-determined genomic and physical distance distribution. The likelihood is taken as the product of these probabilities across both clusters. We systematically transfer each amplicon to the other clusters, and accept the new grouping if the likelihood increases. The cluster assignments are finalized when no increase in likelihood results from transferring any amplicon to another cluster. This approach identified clusters corresponding to territories, as well as smaller clusters corresponding to amplicons that did not colocalize with a territory. Amplicons that did not colocalize with a territory were excluded from downstream analyses.

Aneuploid cells were identified by finding the chromosomes that had more than two clusters for autosomes, or more than one territory for allosomes, and visually confirming that these extra clusters corresponded to territories, and not amplicons unassociated with a territory. These cells were excluded from the final data set (see **Cell Filtering**).

Spatial clustering of reads in embryos
For each chromosome, spatial clusters were identified using k-means clustering (k=2) with semi-supervised correction based on visual inspection of neighboring genomic positions.

When territories for a given chromosome overlapped such that reads could not be confidently assigned to a cluster, the reads were not given a cluster annotation and were specially annotated as overlapping (**Table S3**) (0/475 diploid chromosome pairs in zygotes, 41/779 in 2-cell embryos, 95/872 in 4-cell embryos). For cells that had fewer than 150 total reads (4 cells total, all in the highest z-position of a 4-cell embryo), clusters were not assigned. Clusters were not assigned to Chr Y. During visual inspection, three instances of triploidy were observed for individual chromosomes in 2-cell embryos (embryo 25, cell 1 Chr 10; embryo 25, cell 2, Chr 15; and embryo 44 cell 2 Chr 1).

Following initial clustering, for each chromosome in the embryos, outlier points were detected using a DBSCAN algorithm with a variable search radius of 3.5 multiplied by the mean distance between data points in the nucleus, and the requirement that two neighbor points define a core point. Points that were identified as outliers by the algorithm were excluded from our further analyses.

Calculation of percentage of reads falling outside a chromosome territory
To find the percentage of spatially-localized reads that fell outside of a chromosome territory in PGP1f, the number of reads that were not part of the two largest clusters identified by MLE spatial clustering (or, for allosomes, the single largest cluster) for a chromosome in a cell was divided by the number of reads that were part of the two largest clusters (or, for allosomes, single largest cluster). The resulting percentage was 6.83%.

To calculate an equivalent percentage for embryos , the number of amplicons identified by DBSCAN as outliers in our annotated clusters of chromosomal reads (see *Spatial clustering of reads in embryos*) was divided by the number of amplicons not identified as outliers. The resulting percentage was 6.95%.

Assignment of parent-of-origin to spatially-localized reads
Genotype information (.vcf files) for the B6C3F1 and B6D2F1 mouse strains was obtained from https://www.sanger.ac.uk/data/mouse-genomes-project/. All spatially-localized reads overlapping a SNP unique to one of the two strains were identified. If a read contained a non-reference base at a heterozygous SNP position, it was assigned to the corresponding parent-of-origin. However, if the non-reference base was present in less than 90% of a read's PCR duplicates, it was excluded as a putative sequencing error. Any reads containing SNPs from both the maternal and paternal genomes were marked as conflicting.

Assignment of parent-of-origin to territories
For embryos, parent-of-origin was assigned to chromosome territories when possible. For zygotes, parent-of-origin was assigned manually, by visually identifying which pronucleus each territory belonged to and assigning paternal origin to the territories in the larger pronucleus. For 2- and 4-cell embryos, the following procedure was used: 1) assign each territory to a parent-of-origin based on the majority of parent-of-origin-specific SNPs in the territory, leaving the territory unassigned if it did not contain any SNPs or a majority could not be determined. For cells containing only one territory for the X-chromosome, assign it to maternal origin; 2) for territories left unassigned, if their homolog was assigned, then assign them to the opposite assignment of their homolog; 3) if the parent-of-origin

assignments for two homologs are the same, but one of the assignments was based on a 1-0 majority in parent-of-origin SNPs and the other was not, resolve the contradiction by changing the homolog with the 1-0 majority; 4) if a contradiction in parent-of-origin assignment between homologs could not be resolved, drop both assignments. SNPs in outlier reads were not used in the assignments. Chromosomes that were aneuploid or whose homologs spatially overlapped in a given cell were excluded from parent-of-origin assignments. This procedure led to parent-of-origin assignments for 75.6% of territories in 2-cell embryos and 43.7% of territories in 4-cell embryos. To validate this method of assigning parent-of-origin, we predicted parent-of-origin for zygotic chromosome territories and compared the assignments with ground-truth manual assignments. We found that 79.6% of zygotic chromosome territories could be assigned by this method, of which 97.1% agreed with manual assignments.

Relative radial distance from nuclear center
For PGP1f, to compare the radial positions of chromosomes, we first constructed a 2D convex hull for each cell using the 2D projection of each read within the cell. We then found the relative radial position of each read in the cell. To do this, we defined r_0 as the vector connecting the centroid of the hull and a given read, and then, defined a positive constant c as the minimum scaling factor required for r_0 to intersect a facet of the convex hull. We minimized this constant by solving the equation of the plane for the line coincident with r_0 and each facet of the convex hull. 1/c is thus the relative radial distance of the read to the nuclear center. The radial position for each chromosome was defined as the mean radial position of all reads mapping to that chromosome, aggregated across all cells. 95% CI was determined by bootstrapping.

Repetitive DNA element annotation
For PGP1f, both uniquely-mapping and multi-mapping spatially-resolved reads were annotated for overlap with repetitive DNA elements. For uniquely-mapping reads, the genomic positions of each paired-end read were looked up from the BAM file and examined for overlap with the RepeatMasker database (*40*) using bedtools (*99*). For multi-mapping reads, this process was performed for all potential genomic alignments, and the resulting annotations were collapsed into a single entry.

PGP1f analysis

Genome-wide mean distance map
To construct a genome-wide distance map, we first constructed a square matrix with a size equal to the linear hg38 human genome binned at 10 Mb. Then, for each diploid-resolved chromosome territory, we computed the pairwise distance between each read in the territory and every other read in its home cell *except* for the reads associated with its homologous pair in that cell; we then assigned each pairwise distance to its corresponding bin in the matrix. Finally, we flattened the matrix by computing the mean of each bin.

Comparison of repetitive element frequency
All repetitive element (rep_name, rep_family, and rep_class) annotations from the RepeatMasker database (40) were used for this analysis. To calculate the frequency of each

repetitive element in IGS data, the number of reads corresponding to each element was divided by the total number of spatially-resolved reads. To calculate the approximate frequency of each repetitive element in the reference genome, the number of genomic bases corresponding to each element in the RepeatMasker database (40) was summed and divided by 3 billion.

Radial distribution of repetitive elements
All repetitive element (rep_name, rep_family, and rep_class) annotations from the RepeatMasker database (40) with a frequency greater than 50 were used for this analysis. A distribution for each repetitive element was created by aggregating the relative radial distance measurements for that element across all cells. The repetitive element annotations were then shuffled for each cell 500 times, and the same aggregation was performed on each permutation to create a null distribution for each element. The observed and null distributions for each element were reorganized into 100 bins, each representing a ring 0.05-0.1 microns wide from the nuclear center. A z-score was calculated for each bin and each element, with negative values indicating depletion, and positive values indicating enrichment. The z-scores were then converted to log10(p-values) for visualization. The repetitive elements displaying the most radial bias were identified by selecting those with the greatest variability in bin values. The identified elements with the strongest radial bias were ordered by enrichment profile by sorting by the position of bin with the highest value.

Relationship between genomic and spatial distance
To characterize the relationship between genomic and spatial distances for ensembles of chromosomes in PGP1, we first computed the pairwise genomic and spatial distances between each read localized within each individual diploid-resolved chromosome territory, separately keeping track of whether each measurement was within an arm or between arms. For visualization, these measurements were binned (typically at 1 or 3 Mb) either together (**Fig. S15, S16A-B**), or separately (e.g. **Fig. 2, S16B-C, S17, S18**), and were plotted as bin means ± SD. Power laws were fit to mean pairwise distance measurements (binned at 1Mb) by nonlinear least squares, and 95% confidence intervals were estimated from the sample variance. Residuals were calculated from the bin means and the power law fit and tested for autocorrelation using the Ljung-Box test at lag 1. To compare the distributions of intra- and inter-arm distances, bins were filtered such that only genomic distances containing at least 20 of both types of measurement were retained, denoting a range of shared genomic distances. The spatial distance distributions of intra-arm and inter-arm measurements in this range were then compared by KS test.

In embryos, we took a similar approach, with a few differences. First, reads were distinguished by parent-of-origin rather than homolog. Second, curves for individual pronuclei were also computed.

Developmental transitions analysis

Separation score
Separation scores were calculated for each spatially-localized read by taking its 100 nearest neighbors in 3D space (excluding reads belonging to the same chromosome territory) and

dividing the number of neighbors assigned to either the maternal or paternal genome by the total number of neighbors with an assignment, and selecting the higher fraction. For the boxplots shown in **Fig. 4C**, a mean separation score for each nucleus was calculated by averaging the separation score of all its spatially-localized reads.

Centromere polarization
Nuclei from the embryo data set were partitioned by developmental stage. To quantify the degree of centromere polarization at each stage, the spatially-localized reads closest to each segmented CENP-A loci were selected. The spatial positions of these reads were averaged to create a weighted center, and then the distance from the weighted center to the center of the entire nucleus was measured. This process was repeated for each nucleus. To determine if the distribution of distances between the weighted centromere centers and nuclei centers was significant, the distances to the centromere in each nucleus were randomly permuted. The same weighted center analysis was performed, and the significance for each stage was calculated by performing a two-sample K-S test on the real and permuted distributions.

Rabl-like configuration
Each spatially-localized read was assigned a centromere-telomere score between 0 and 1 by dividing its genomic position by the total length of the corresponding chromosome. Rabl scores were calculated for each read by taking its 100 nearest neighbors in 3D space (excluding reads belonging to the same chromosome territory) and averaging the centromere-telomere score of its neighbors. For the plot shown in **Fig. 4E**, reads were first partitioned by developmental stage and then split into 100 bins according to their centromere-telomere score (from 0-0.01 to 0.99-1). The "mean neighbor telomere-centromere position" shown on the y-axis was calculated by taking the median Rabl score for all reads falling within a particular bin. The r values reported represent the Pearson correlation between all spatially-localized reads' centromere-telomere scores and Rabl scores.

Relationship between GC content and distance to nuclear landmarks
Spatially-localized reads in the zygote and 2-cell stage were partitioned into non-overlapping haplotype-resolved 1 Mb bins spanning mm10. The GC content of each bin was calculated using the mm10.gc5Base.bw file from the UCSC Genome Browser. The average distance of each bin to the nuclear lamina and to the nucleolus precursor bodies (NPBs) was calculated by iteratively averaging all of the reads in the bin. The distances from reads originating from the same cell were averaged together first, and then all remaining distances were averaged yielding a final value. Correlations between GC content and average distance to nuclear landmarks were performed for each homolog separately and were reported as Spearman's ρ.

Single-cell domain analysis

IGS comparison with Hi-C
Allele-resolved Hi-C data for zygotes and 2-cell embryos were obtained from a published study (*48*). allValidPairs.txt.gz files were temporarily split, lifted over from mm9 to mm10 using the UCSC liftOver utility, and re-combined. The re-combined files were then

converted to .hic format (*100*) and loaded into Juicer (*101*) to calculate compartment eigenvalues for each 1 Mb bin in mm10.

Next, using our IGS data, a lamin proximity score was calculated for each 1 Mb bin in the mouse genome, defined per-bin as the mean-centered probability that a read was closer than 500 nm to a lamin immunostain. The Pearson correlation coefficients of the Hi-C eigenvalues and the lamin proximity scores were determined per chromosome.

Single-cell domain boundary detection
To detect chromatin domains in single cells, first, a mean pairwise distance matrix was constructed at 2.5 Mb resolution for each single chromosome in each single paternal pronucleus. Matrices were then filtered using a 90% coverage threshold. For the remaining high-coverage matrices, gaps were resolved by linear interpolation. To detect and score boundaries, we adapted an approach used in (*74*). Within a sliding window along the diagonal, we calculated all pairwise distances between bins on each side of the window center ('intra-domain') and between bins on opposite sides of the center ('inter-domain'). The mean intra-domain and inter-domain distances were then calculated and an insulation score was assigned at the window center, defined as the difference of the means divided by the sum of the means. SCD boundaries were then defined as local maxima in the resulting insulation score vector, with peak-finding parameters set by visual inspection of boundary calls in representative cells. To avoid edge noise, peaks found within three bins of the start or end of the matrix were not considered. Boundaries and insulation scores were also called and calculated for ensemble matrices using the approach described above.

Scaled domain distance from lamin
To examine the relationship between SCDs and the nuclear lamina, first, domain boundaries were found in high-coverage matrices as described above. Next, the 'observed distance-to-lamin profile' for scaled domains was constructed: for each sequential pair of boundaries, we considered all reads from that single chromosome with a genomic position falling between the boundaries. Using a fixed number of bins (N) for all pairs of boundaries in all chromosomes, each read was binned based on its relative genomic position between the two boundaries. The distance to the nuclear lamina for that read was then recorded in the corresponding bin. After processing all reads, the rightmost N/2 bins of distances were reflected and vertically concatenated to the leftmost N/2 bins of distances, resulting in a final observed distance-to-lamin distribution spanning a scaled distance of 0 to 0.5 (i.e. 0 to N/2) from SCD boundaries. To calculate the expected distance-to-lamin vector, for each pair of observed boundaries, the position of the boundary-pair was randomly shifted within the chromosome, while keeping the genomic distance between the two boundaries constant (to respect the chromosome edges shifted boundaries were required to remain within the minimum and maximum matrix boundaries). The read positions and lamin-distances were then evaluated and binned for the shifted boundaries as described above. This approach was used to generate a number of random samples for each chromosome proportional to the genomic size of the chromosome, and the final 'expected distance-to-lamin profile' was constructed as described above. A 95% confidence interval for the median observed-over-expected distance from lamin was then determined per bin by bootstrapping.

Global chromosome positioning analysis

Construction of single-cell autosome distance matrices
For each autosome territory, the mean spatial position of all reads in the territory was calculated to find approximate centers for the territories. For homolog-resolved but non-haplotyped analysis (**Fig. 6**, **Fig. 26A**), a 19 x 19 single-cell autosome distance matrix was constructed by 1) for each pair of chromosomes, finding the distances between the centers of all inter-pair homologs (ie between Chr A homolog 1 and Chr B homolog 1, Chr A homolog 1 and Chr B homolog 2, etc, without regard for parent-of-origin); and 2) taking the mean inter-pair distance. This procedure allowed construction of a distance matrix that accounted for the separate positioning of each chromosome territory while remaining agnostic to haplotype, allowing the use of all chromosome territories regardless of whether haplotype could be assigned. For chromosomes in which the two territories were overlapping, and thus could not be confidently broken into two clusters (see *Spatial clustering of reads in embryos*), reads were assigned randomly into two clusters, since their pairwise distances with other chromosomes would be very similar due to their spatial co-localization. For the three cells with instances of triploid chromosomes (see *Spatial clustering of reads in embryos*) (**Fig. S19**), those chromosomes were excluded from analysis, resulting in blank entries (NaN) in columns and rows corresponding to the triploid chromosome in the distance matrix for that cell.

For haplotype-resolved analysis (**Fig. S26B**), the single-cell autosome distance matrix was constructed by finding the pairwise distance between all autosome territories for which parent-of-origin was assigned, leaving blank entries (NaN) in columns and rows corresponding to autosome territories for which parent-of-origin could not be assigned.

Correlation of single-cell autosome distance matrices
To find the correlation between two single-cell autosome distance matrices, we first unraveled the upper triangular of each distance matrix, including the diagonal for haplotype-agnostic analysis and excluding it for haplotype-resolved analysis, and then calculated the Pearson correlation coefficient between the two vectors. For haplotype-resolved analysis and cells which contained triploid chromosomes, the single-cell autosome distance matrices had missing entries (see above), so the vectors were modified to contain only entries that were present in the matrices for both cells being compared. Pairs of cells that shared fewer than 3 sets of haplotyped autosomes (6 territories) were discarded from the haplotype-resolved analysis.

Construction of putative clonal lineage trees
For each 4-cell embryo, the most correlated pair of cells was taken to be a pair of putative sister cells. When all four cells in the embryo passed the threshold of 150 reads per cell (see *Spatial clustering of reads in embryos*), the other two cells were taken to be the second pair of putative sister cells, which was also the second-most correlated pair in 5/6 such embryos (**Fig S27**). All other pairs were taken to be putative cousin cells. When only three cells in a 4-cell embryo were available, the pairs of cells that were not the most correlated were taken to be putative cousin cells.

Supplementary Text

Kit-free synthesis of in situ sequencing reagents

We note that while we purchased the SOLiD sequencing reagents commercially, they are no longer being sold as a kit. However, while we were preparing our manuscript, Nguyen et al. (72) reported methods for producing reagents functionally equivalent to SOLiD sequencing reagents independently of Applied Biosystems. We provide this note as guidance to those who wish to produce their own sequencing reagents.

SOLiD Sequencing Oligos
Nguyen et al. recently reported that oligos containing the cleavable backbone linker used in SOLiD sequencing oligos can be obtained from Integrated DNA Technologies (IDT) using internal 3'-thio deoxyinosine, an off-catalog internal oligo modification that can be requested as a custom order, followed by a phosphorothioate modification. The IDT key for the modification is N/i3Thio-dI/N, where N is any nucleotide, and /i3Thio-dI/ is internal 3'-thio deoxyInosine.

Based on this report in Nguyen et al., we recommend the oligonucleotide sequences in **Table S5,** which are functionally equivalent to SOLiD oligos.

Sequencing Oligo Cleavage Buffers
Nguyen et al reported that 50 mM AgNO3 in H2O can be used as a replacement for SOLiD Buffer C. We have found that the buffer described in **Table S6** can be used as a replacement for SOLiD Buffer B.

Cost, complexity, and throughput

Cost
Table S7 provides a cost breakdown estimate for the most costly reagents required to perform IGS at a scale equivalent to our experiment with early mouse embryos. Volumes correspond to library preparation and in situ sequencing of a pooled experiment performed using 300 μl total volume. We note that reagent cost is dominated by ex situ Illumina sequencing, which could be reduced by sequencing fewer reads. In our sequencing data generated by NovaSeq, we obtained a duplication rate of 99.72%, indicating library saturation and the potential for a reduction in ex situ sequencing.

Complexity
In addition to the reagents described in **Table S7**, IGS requires a dedicated confocal fluorescence microscope for overnight automated imaging in order to collect many bases of in situ sequencing data. Imaging requires setup of automated imaging protocols, which is accessible within imaging software packages (we used NIS-Elements AR and Andor Fusion 2.0 imaging software to automate imaging of PGP1f and early mouse embryos, respectively, as described in the *Materials and Methods* section).

For automated sequencing, which we used in the collection of data from PGP1f, an automated fluidics setup is required. As described in the *Materials and Methods* section, we performed automated fluidics by using a custom MATLAB script to control a modular valve positioner and peristaltic pump, which were connected to a computer via a National Instruments Data Acquisition card.

Throughput
The throughput-limiting step for IGS is imaging time for in situ sequencing, which scales linearly with the number of bases sequenced, the number of fields of view, the thickness of the sample, and the imaging exposure time. Using the exposure times, optimized imaging buffer, and z-step size we used for imaging in embryos, we estimate that 49.4 cells, 1.2 embryos, or simply 123.5 z-steps could be collected per hour of total imaging time over 18 bases of imaging. To estimate the times per embryo and per cell, we have assumed the average z-height required to capture IGS data from embryos in our dataset (41.7 µm), cells with 10 µm nuclei, and 10 cells per FOV.

We note that speed could be improved by using brighter, more photostable dyes for the sequencing reagents. This is an immediate possibility using custom synthesis of the sequencing oligos, as described in the section *Kit-free synthesis of in situ sequencing reagents* in the Supplementary Text.

Supplementary Figures

DAPI after fixation
and permeabilization

DAPI after
library preparation

Library preparation
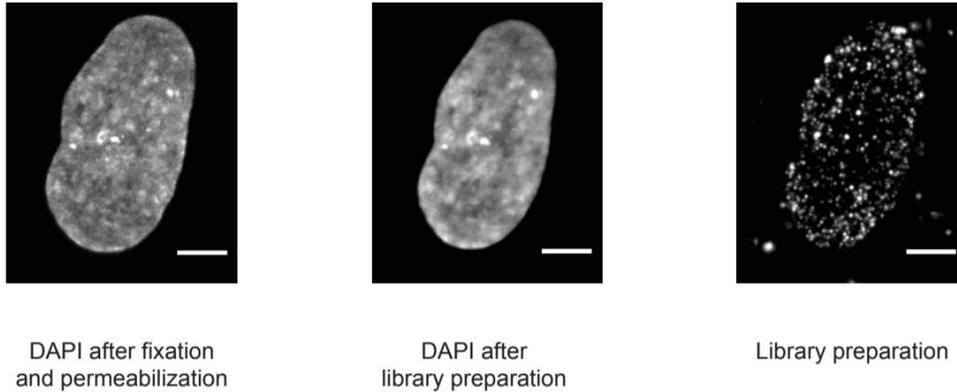
Fig. S1
DAPI staining before and after library preparation. DAPI staining in a PGP1f nucleus after fixation and permeabilization (left) and after library preparation (middle) shows that morphological features are well-preserved during library preparation. Slight shrinkage of nuclei was observed over the course of library preparation, in line with what is seen during DNA FISH protocols (*89*). (Right) IGS library in the same nucleus, stained with amplicon visualization oligo. Scale bars, 5 μm.

Library

Primer 1 Primer 2 Primer 3 Primer 4
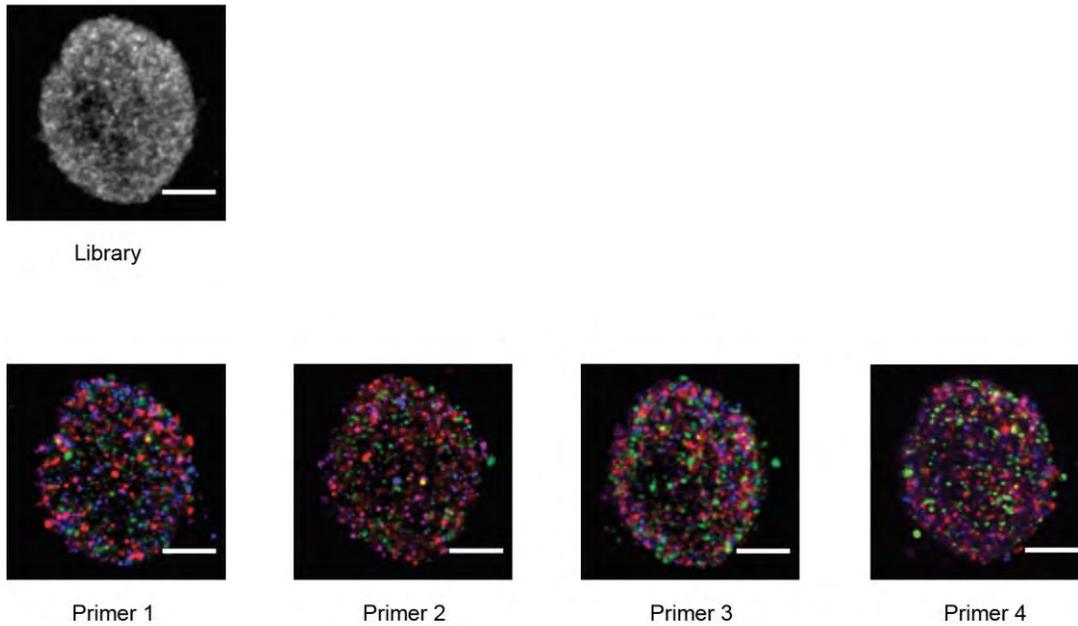
Fig. S2
Sub-sampled sequencing of a high-yield library. (Above) A single z-plane through a representative PGP1f nucleus in a high-yield *in situ* genomic sequencing library stained with amplicon visualization oligo. (Below) One base of *in situ* sequencing in the same nucleus using each of four orthogonal *in situ* sequencing primers. The effective density of high-yield libraries can be modulated by using a subset of these primers. All amplicons are amplified for NGS analysis regardless of subsampling, and PCR amplicons generated from unimaged *in situ* amplicons are discarded during the UMI matching step. Scale bars, 5 μm.
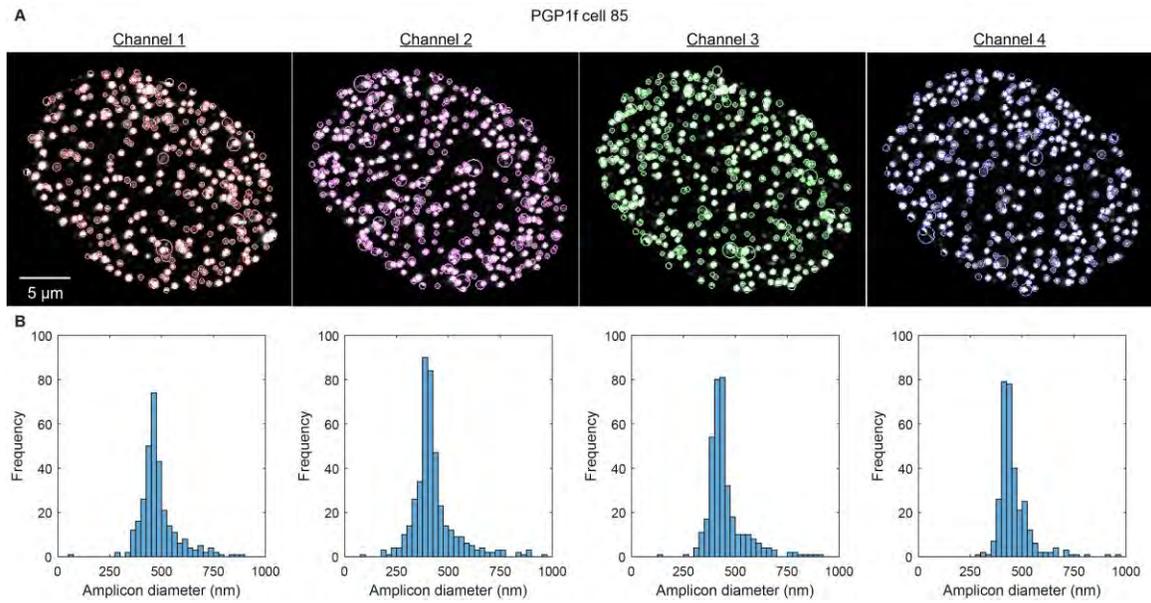
Fig. S3
Quantification of amplicon size. (**A**) Representative nucleus (PGP1f cell 85) across all four sequencing channels with superimposed circles of radius 2σ (where σ is the standard deviation of the Gaussian fit for each amplicon). The images are maximum intensity projections across z-planes for the first cycle of *in situ* sequencing. (**B**) Histograms of the amplicon diameters (4σ) measured from each channel.
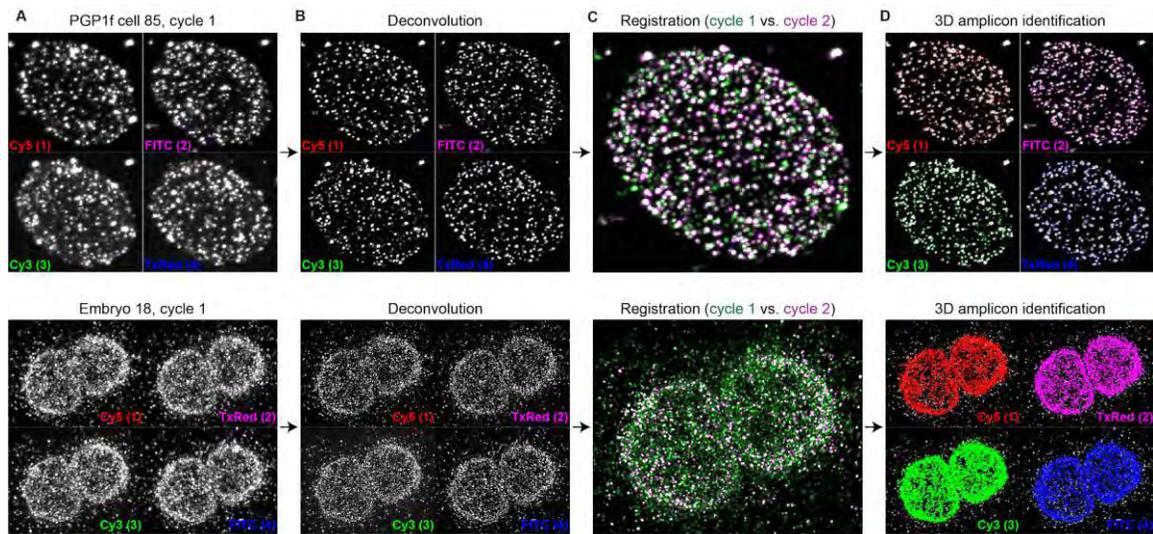
Fig. S4

Nuclei image processing. (**A**) Raw *in situ* sequencing images (cycle 1) split by imaging channel for PGP1f cell 85 (top) and Embryo 18 (bottom). Preprocessing steps are performed in 3D; all displayed images are maximum intensity projections across channels and z-planes (all z planes for PGP1f, z = 50-70 = 5 μm slice for the embryo) for visualization purposes. (**B**) Deconvolution of *in situ* sequencing images (cycle 1) split by imaging channel for PGP1f cell 85 (top) and Embryo 18 (bottom). (**C**) Registration of *in situ* sequencing images (cycle 1 = green, cycle 2 = magenta) for PGP1f cell 85 (top) and Embryo 18 (bottom). (**D**) 3D amplicon identification from *in situ* sequencing images (cycle 1) split by imaging channel for PGP1f cell 85 (top) and Embryo 18 (bottom). Colored circles represent identified amplicons by imaging channel.
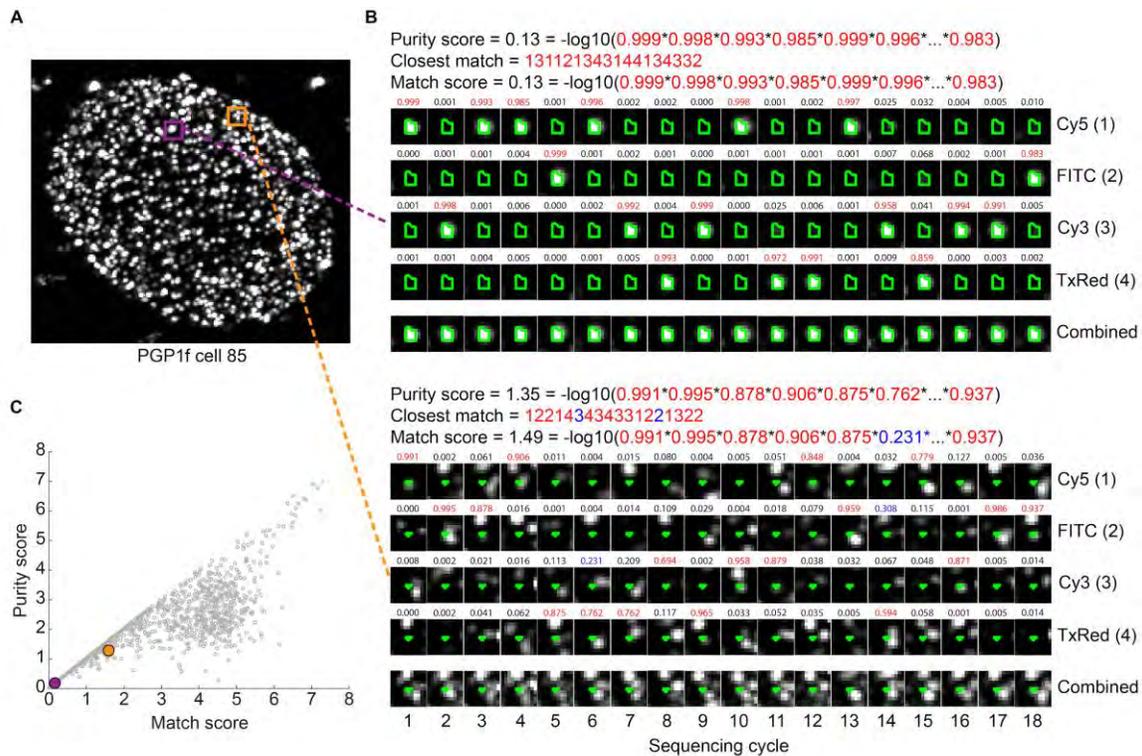
Fig. S5

Amplicon quantification and UMI matching. (**A**) Representative nucleus (PGP1f cell 85) with two spatially-resolved amplicons highlighted. Image is a maximum intensity projection across channels and z-planes for the first cycle of *in situ* sequencing. (**B**) Quantification and UMI matching of the two amplicons from (A). The top set of images corresponds to the amplicon boxed in purple (in a sparsely-packed region), while the bottom set of images corresponds to the ampicon boxed in orange (in a densely-packed region). The bottom row of each set of images is a maximum intensity projection of the four channels, useful for visualizing amplicon density in the region. The green outline in each image indicates the region being quantified, while the number above each image represents the percentage of cycle fluorescence found in the corresponding channel (the sum of all channels in each cycle = 1). The purity score for each amplicon is calculated by multiplying the highest percentages for each cycle (indicated in red) and taking the negative log transformation of the product, i.e. -log10(product(max(matrix,1),2)). Each amplicon also has an associated closest UMI match. While the maximum channels of the top amplicon perfectly match its closest UMI, the bottom amplicon has two positions where the closest UMI match doesn't correspond to the maximum channels (indicated in blue). Taking the percentages from this path to calculate a match score results in a value that is greater (i.e. worse) than the lower bound of the purity score. (**C**) Comparison of purity score and match score for all amplicons in the representative nucleus shown from (A). The points corresponding to the two amplicons highlighted in (B) are enlarged and color-coded.
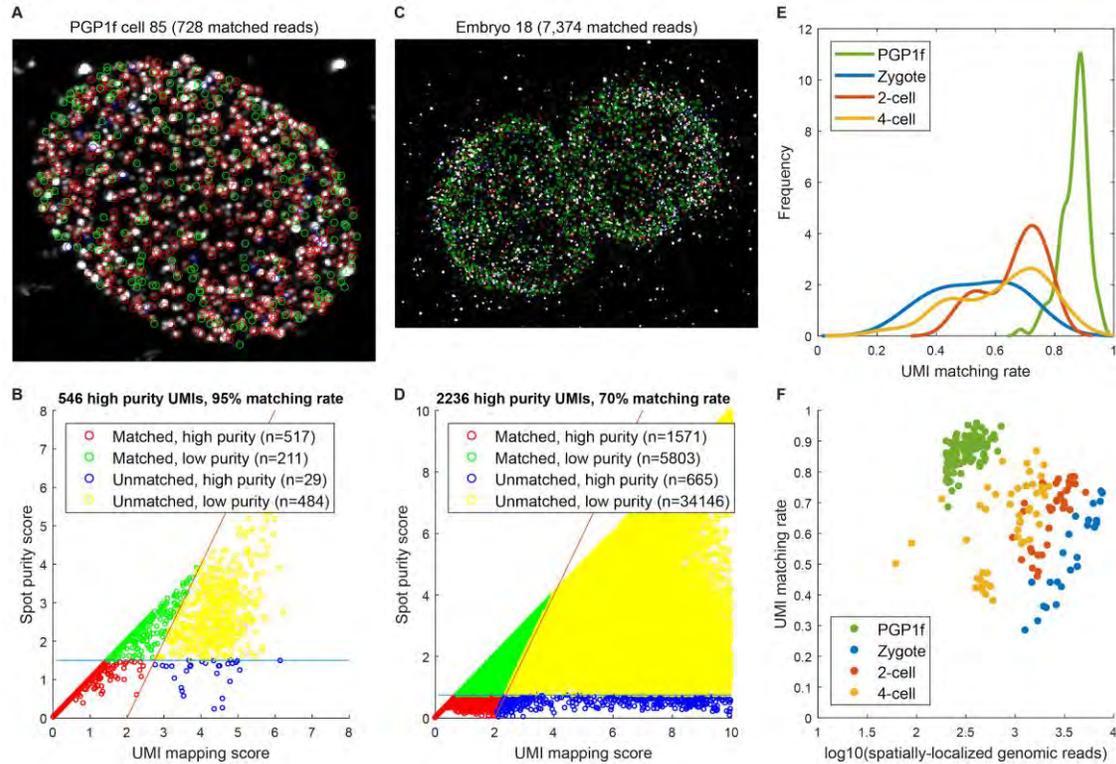
Fig. S6

UMI matching rate. (**A**) Overlay of UMI matching status and a maximum intensity projection across all channels and z-planes for the first cycle of *in situ* sequencing for PGP1f cell 85. Amplicons are colored by UMI matching status, see legend in (**B**). Labels for unmatched, low purity amplicons are omitted for clarity. (**B**) All amplicons from PGP1f cell 85 plotted by match score and purity score. The horizontal line represents the threshold for high purity amplicons (1.5); the diagonal line represents the continuous threshold used to select valid UMI matches. (**C**) The same as (A), but for a 5 μm slice of Embryo 18 (z = 50 to 70). (**D**) The same as (D), but for Embryo 18. A higher threshold for high purity amplicons (0.75) was chosen based on the higher quality of the embryo *in situ* sequencing images. (**E**) The distribution of UMI matching rate by cell type. (**F**) Relationship between number of spatially-localized genomic reads and UMI matching rate, with points colored by cell type.
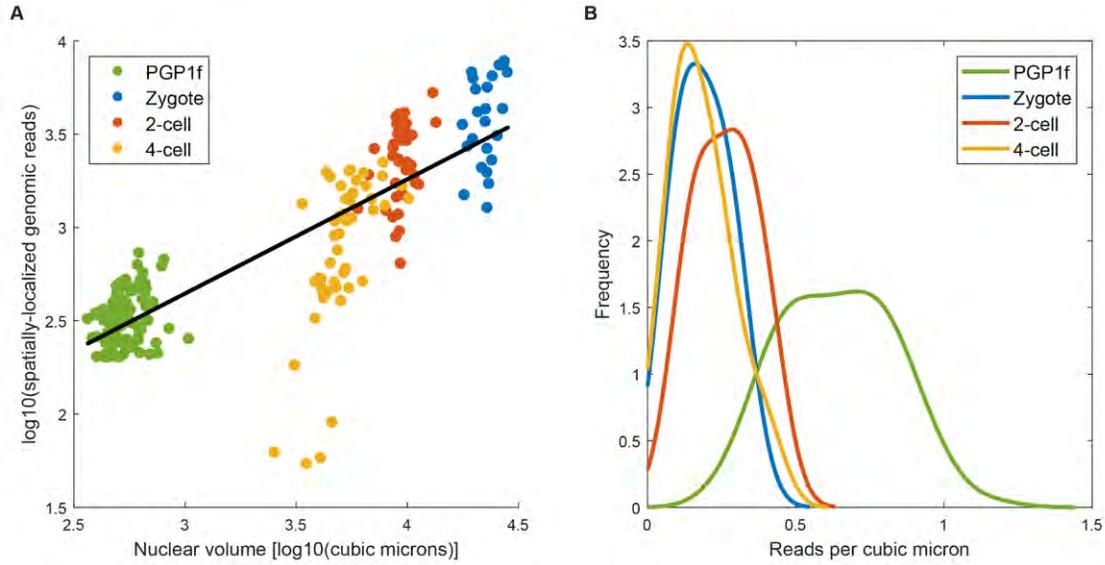
Fig. S7
Relationship between matched reads and nuclear volume. (**A**) Nuclei from PGP1f and all embryonic stages plotted by number of matched reads and 3D nuclear volume, as estimated from DAPI image stacks. The line represents a linear fit on the log-transformed data. (**B**) Distribution of reads per cubic micron by cell-type and embryonic stage. The higher mean in PGP1f is likely attributable to the higher UMI matching rate (see Fig. S6).
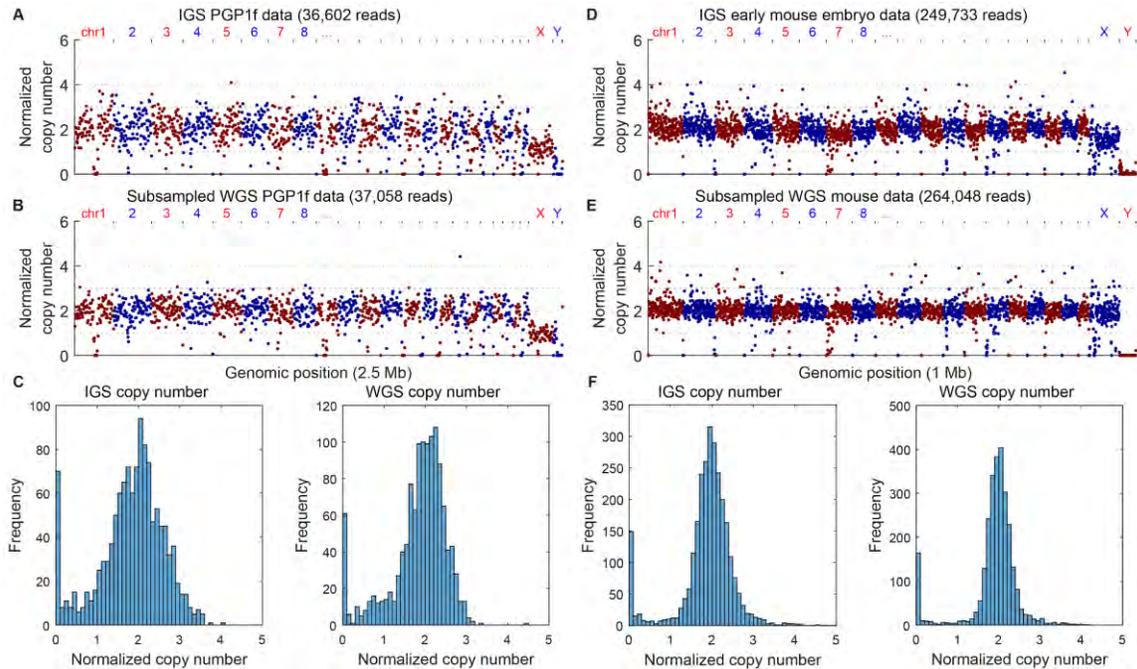
Fig. S8

Comparison of coverage to whole-genome sequencing. (**A**) Normalized coverage of PGP1f IGS data across 2.5 Mb bins in hg38 (n=36,602 reads). (**B**) Normalized coverage of subsampled PGP1f whole-genome sequencing data (ENCODE accession ENCFF713HUF) across 2.5 Mb bins in hg38 (n=37,058 reads). (**C**) Histograms showing the distribution of normalized coverage per autosomal bin in IGS (left) and WGS (right). (**D**) Normalized coverage of mouse early embryo IGS data across 1 Mb bins in mm10 (n=248,733 reads). (**E**) Normalized coverage of subsampled mouse whole-genome sequencing data from SRR7511358 across 1 Mb bins in mm10 (n=264,048 reads). (**F**) Histograms showing the distribution of normalized coverage per autosomal bin in IGS (left) and WGS (right).
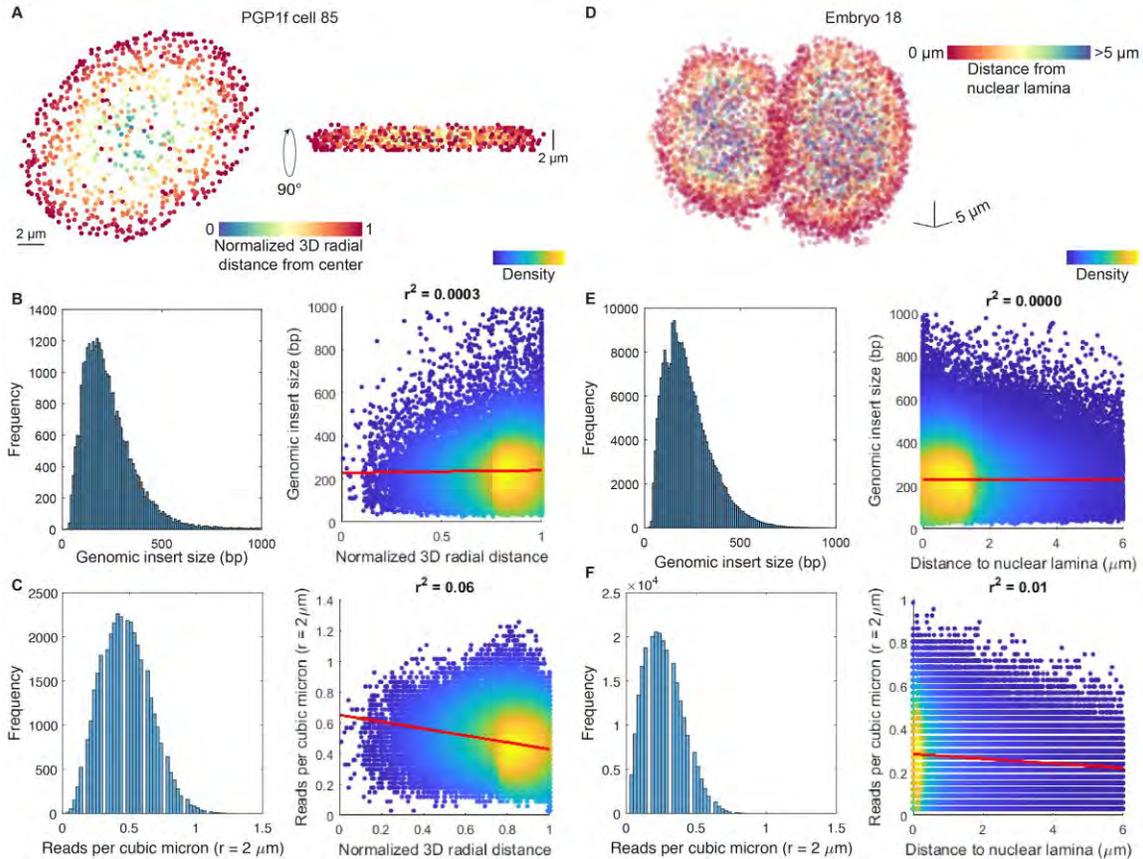
Fig. S9

Fragment size and read density by radial position. (**A**) PGP1 fibroblast (PGP1f cell 85) shown from two angles, 90 degrees apart, colored by normalized 3D radial distance. (**B**) Left: Histogram of genomic insert sizes for all PGP1f reads. Right: All reads plotted by normalized 3D radial distance and genomic insert size, colored by point density. Best fit line shown in red. (**C**) Left: Histogram of reads per cubic micron in a 2 micron radius for all PGP1f reads. Right: All reads plotted by normalized 3D radial distance and reads per cubic micron, colored by point density. Best fit line shown in red. (**D**) Zygote (Embryo 18), colored by distance to nuclear lamina. (**E**) Same as (**B**), but for all reads in embryos. (**F**) Same as (**C**), but for all reads in embryos.
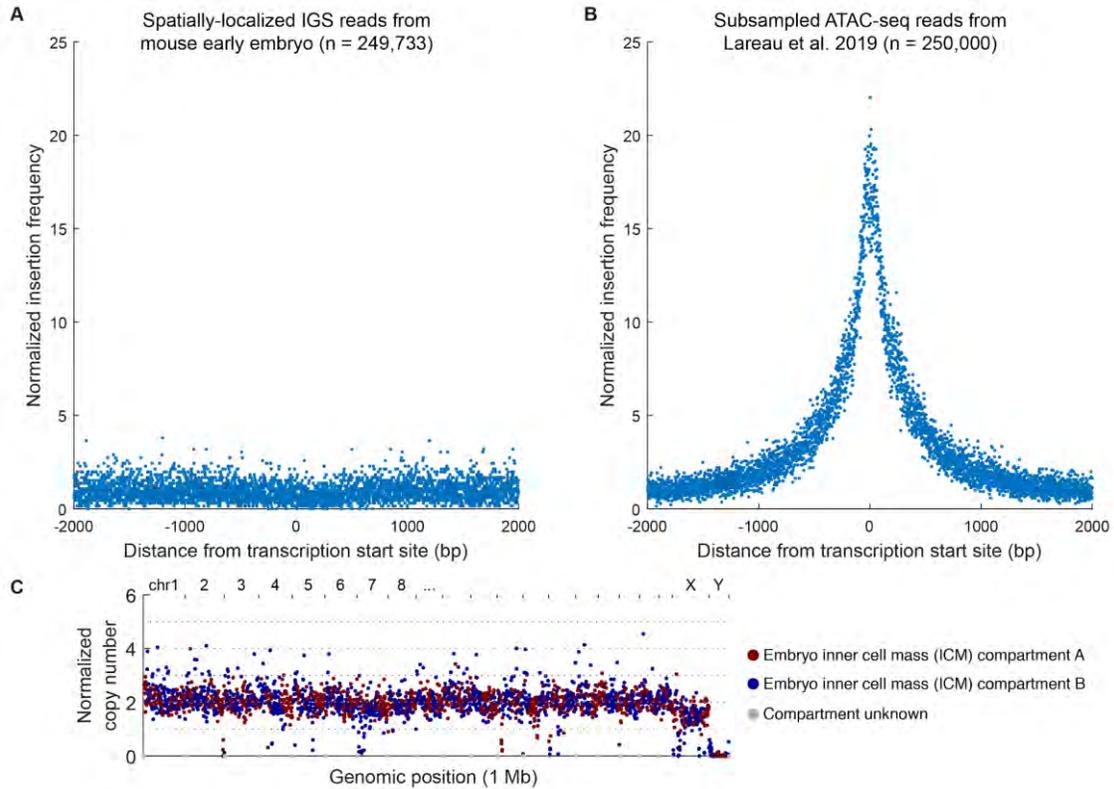
Fig. S10
Lack of enrichment for accessible chromatin. (**A**) Enrichment of genomic reads at genomic positions surrounding transcription start sites in early mouse embryo IGS data (n=249,733 reads). (**B**) Enrichment of genomic reads at genomic positions surrounding transcription start sites in subsampled mouse brain single-cell ATAC-seq data from (*90*) (n=250,000 reads). The difference between these plots indicates that IGS does not have a bias towards accessible chromatin despite similarities in the transposase-based library preparation to ATAC-seq. (C) Normalized coverage of mouse early embryo IGS data across 1 Mb bins in mm10 (n=248,733 reads), colored by embryo inner cell mass (ICM) A/B compartment calls from Hi-C data (*47*).
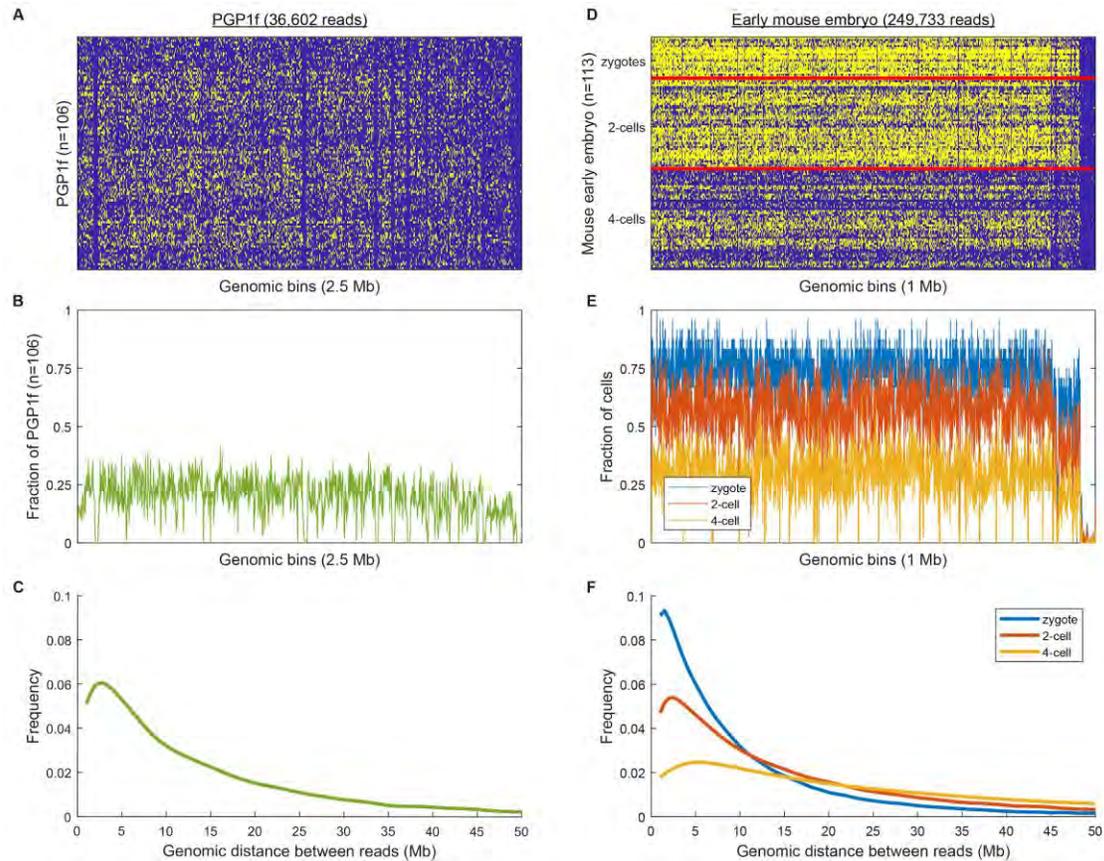
Fig. S11

Sampling of genomic regions in individual cells. (**A**) A *n* by *m* matrix showing which 2.5 Mb genomic bins across hg38 (*m*=1,249) were sampled in individual PGP1f nuclei (*n*=106). Yellow indicates that the genomic region in question was sampled at least once in a given nucleus. (**B**) A line plot showing the fraction of PGP1f nuclei containing at least one read from each 2.5 Mb genomic bin. (**C**) The distribution of genomic distances between sampled loci from the same chromosome territory in PGP1f. (**D**) A *n* by *m* matrix showing which 1 Mb genomic bins across mm10 (*m*=2,737) were sampled in individual early mouse embryo nuclei, ordered by developmental stage (*n*=113). (**E**) A line plot showing the fraction of early mouse embryo nuclei containing at least one read from each 1 Mb genomic bin by developmental stage. (**F**) The distribution of genomic distances between sampled loci from the same chromosome territory in early mouse embryo nuclei by developmental stage.
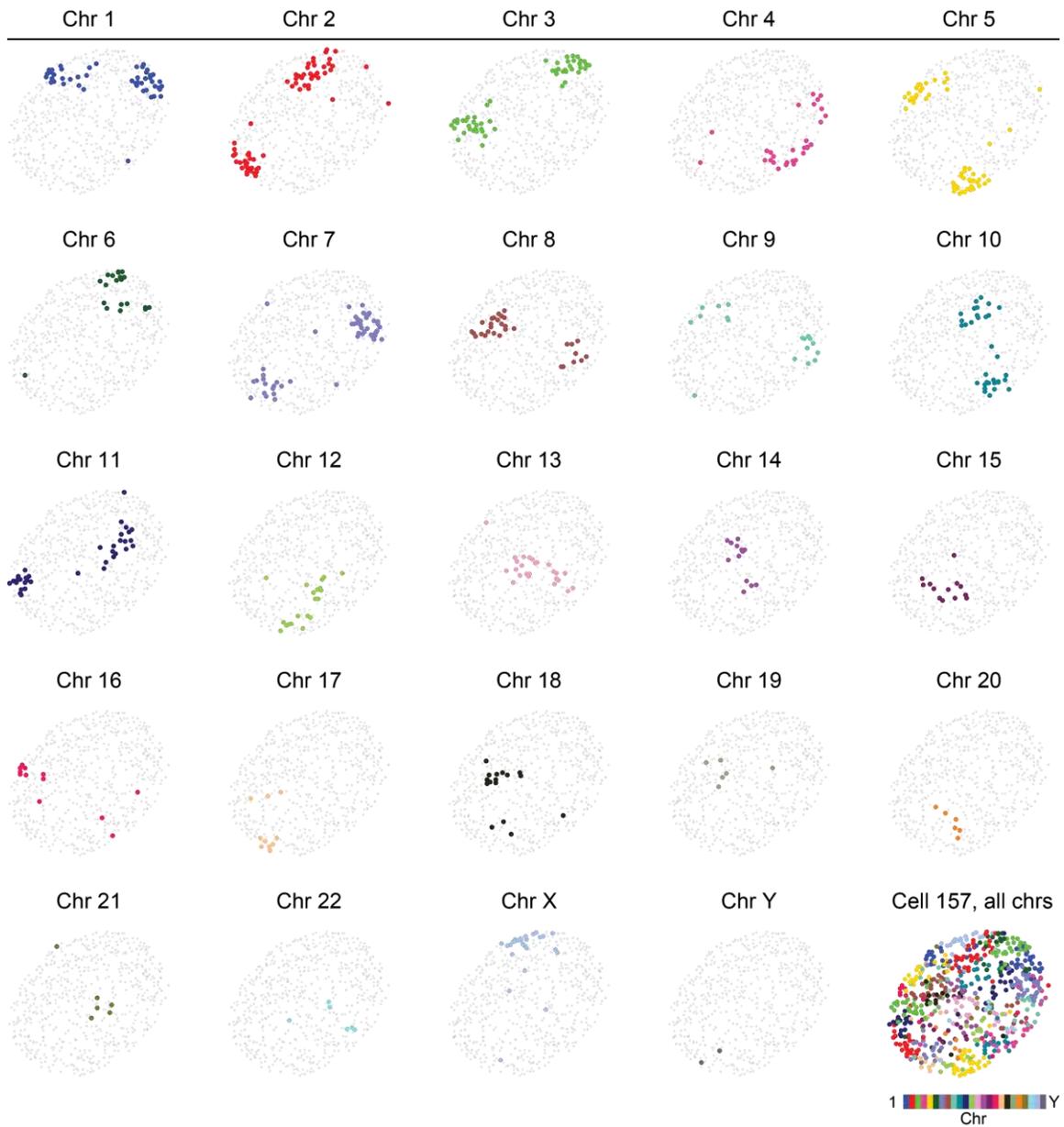
Fig. S12
"2n" colocalization pattern (PGP1f). "2n" co-localization pattern of spatially-localized reads colored by chromosome, before maximum likelihood clustering into territories. Here the measurement is broken down into individual chromosomes in a single cell for visual clarity. Most reads mapping to a given autosome typically cluster into two territories, while reads mapping to sex chromsomes typically cluster into one. ~7% of reads do not colocalize with one of these large territories, possibly due to a combination of UMI FDR and chromosome looping (*4*, *5*).
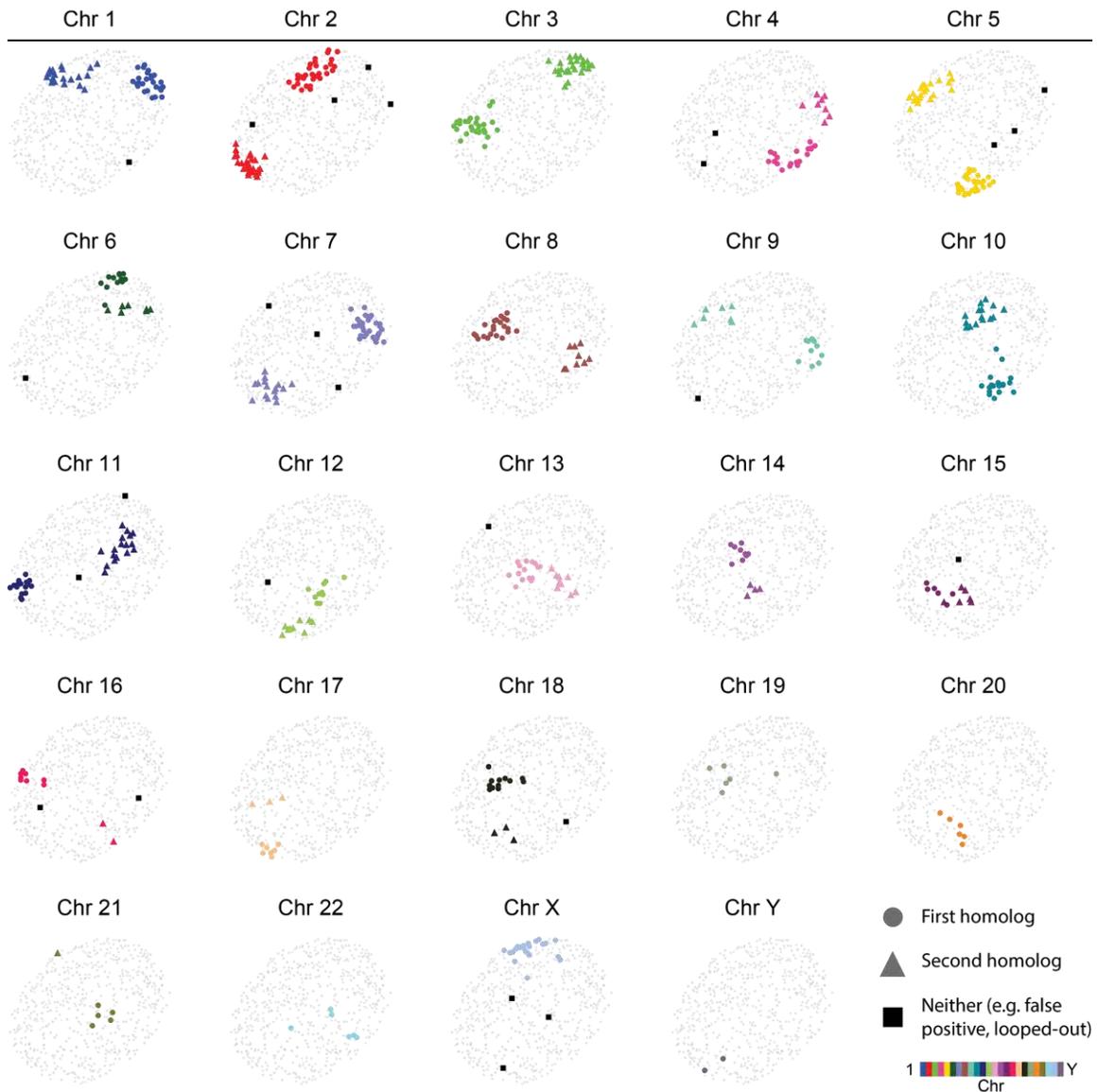
Fig. S13

Territory classification after maximum likelihood clustering (PGP1f). Individual reads in a single cell after maximum likelihood-based clustering, into territories. The discrete spatial nature of individual chromosomes in interphase allows most reads to be assigned to a specific chromosome territory (denoted as a circle or triangle for each homolog), resolving the diploid nature of the genome for these reads. The genomic positions of these reads were used to resolve ambiguities when chromosomes are close to each other (preventing purely spatial classification); nonetheless, small chromosomes are more challenging to resolve than larger ones when they are in close proximity. Reads which do not spatially co-localize with a territory can also be identified by this clustering method, permitting them to be accounted for in downstream analyses.
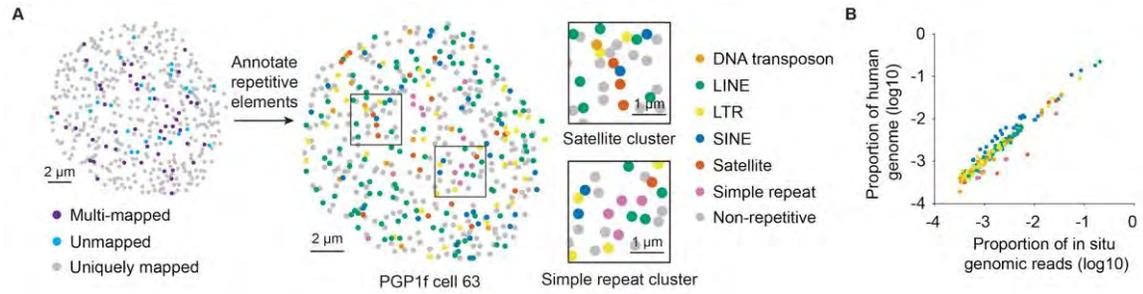
Fig. S14
Repetitive element frequency (PGP1f). (**A**) Spatially-localized reads in a representative nucleus (PGP1f cell 63) colored by alignment status (left) and repetitive element class (right). The boxes highlight two regions that appear to contain spatial clusters of Satellite and Simple Repeat elements respectively. (**B**) The observed number of reads associated with each repeat annotation is proportional to their expected frequency in hg38.
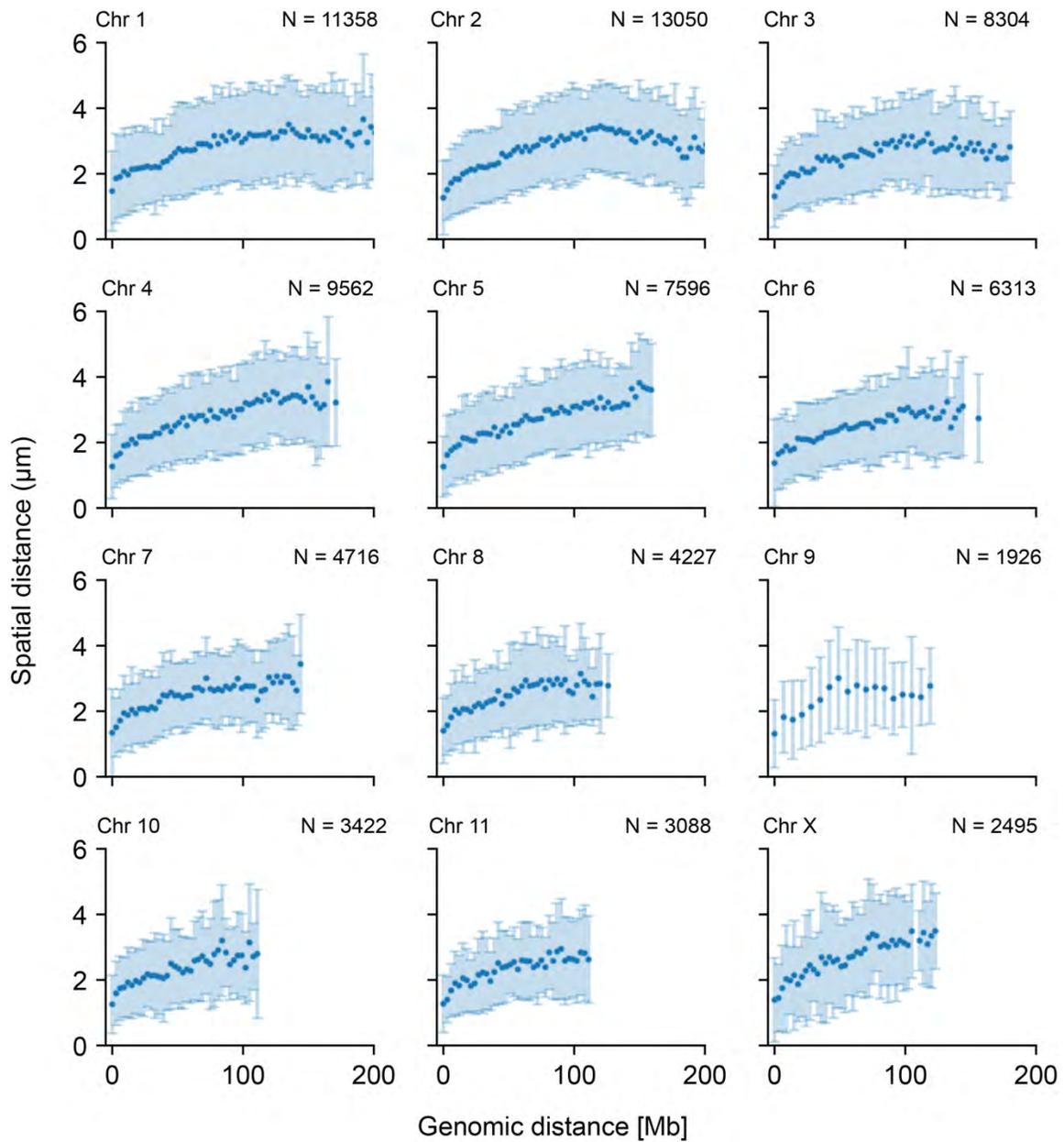
Fig. S15
Relationship between genomic vs spatial distance for Chr 1-11 + X (PGP1f). Distances were computed pairwise for each territory, and binned at 3 Mb (rather than 1 Mb as in Fig. 2I) to ensure sufficient coverage in smaller chromosomes. Chr 9 is an exception and was instead binned at 7 Mb because of a dearth of inter-arm measurements due to its large centromere and small p arm. N, number of pairwise distance measurements. The mean of each bin ± 1 SD is plotted. Bins with fewer than 20 measurements were excluded.
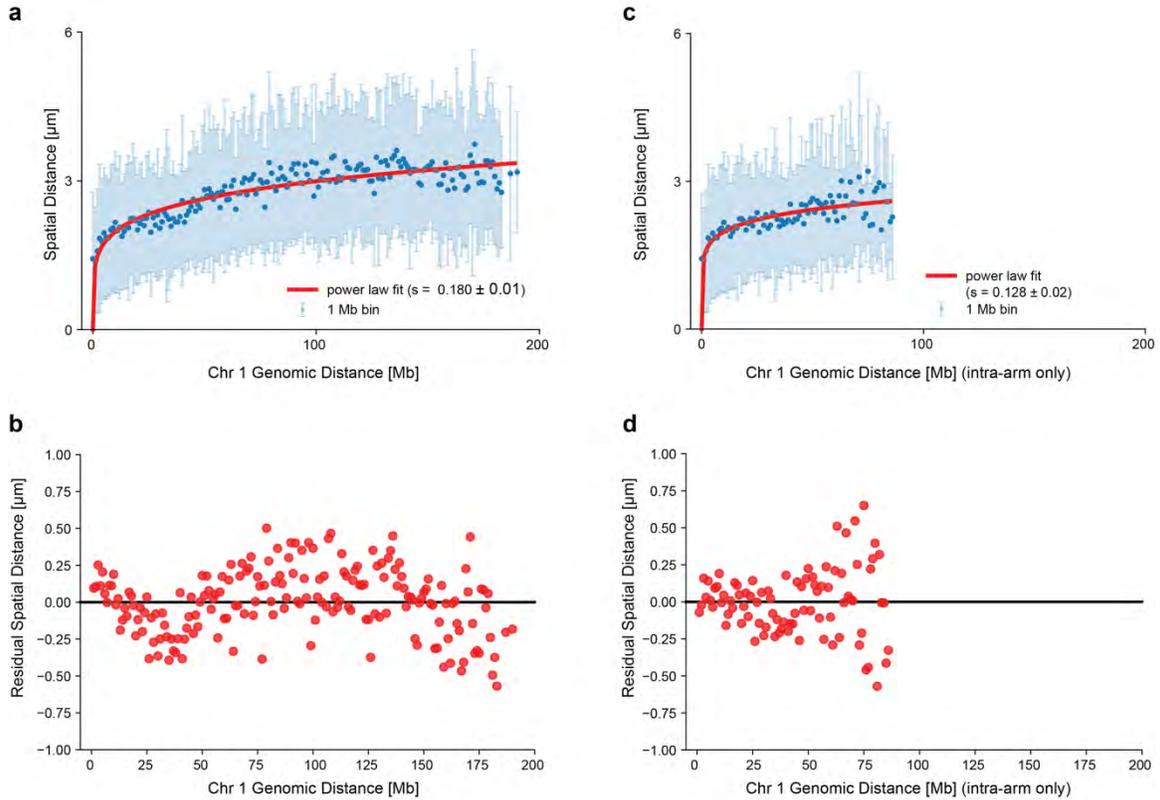
Fig. S16
Chr 1 genomic vs mean spatial distance, power law function fit (PGP1f). (**A**) Ensemble
Chr 1 genomic vs mean spatial distance. Distances were computed pairwise for each Chr
1 territory (n = 144 territories, 11358 pairwise distances), binned at 1 Mb, and plotted
(mean ± SD) alongside a power law function fit to these measurements (scaling exponent
s = 0.180 ± 0.01, 95% CI). (**B**) Residuals from the power-law fit to all of Chr 1 in a). The
residuals are nonuniform around 0 and are autocorrelated at lag 1 (Ljung-Box test, $p < 10^{-5}$). (**C**) Pairwise distances restricted to intra-arm measurements (n = 5788 pairwise
distances), yielding a smaller scaling exponent (s = 0.128 ± 0.02). (**D**) Residuals from the
intra-arm power law fit in (C). The residuals are not significantly autocorrelated at lag 1
(Ljung-Box test, p = 0.46). For both (A) and (C)**,** bins with fewer than 20 measurements
were excluded.

37
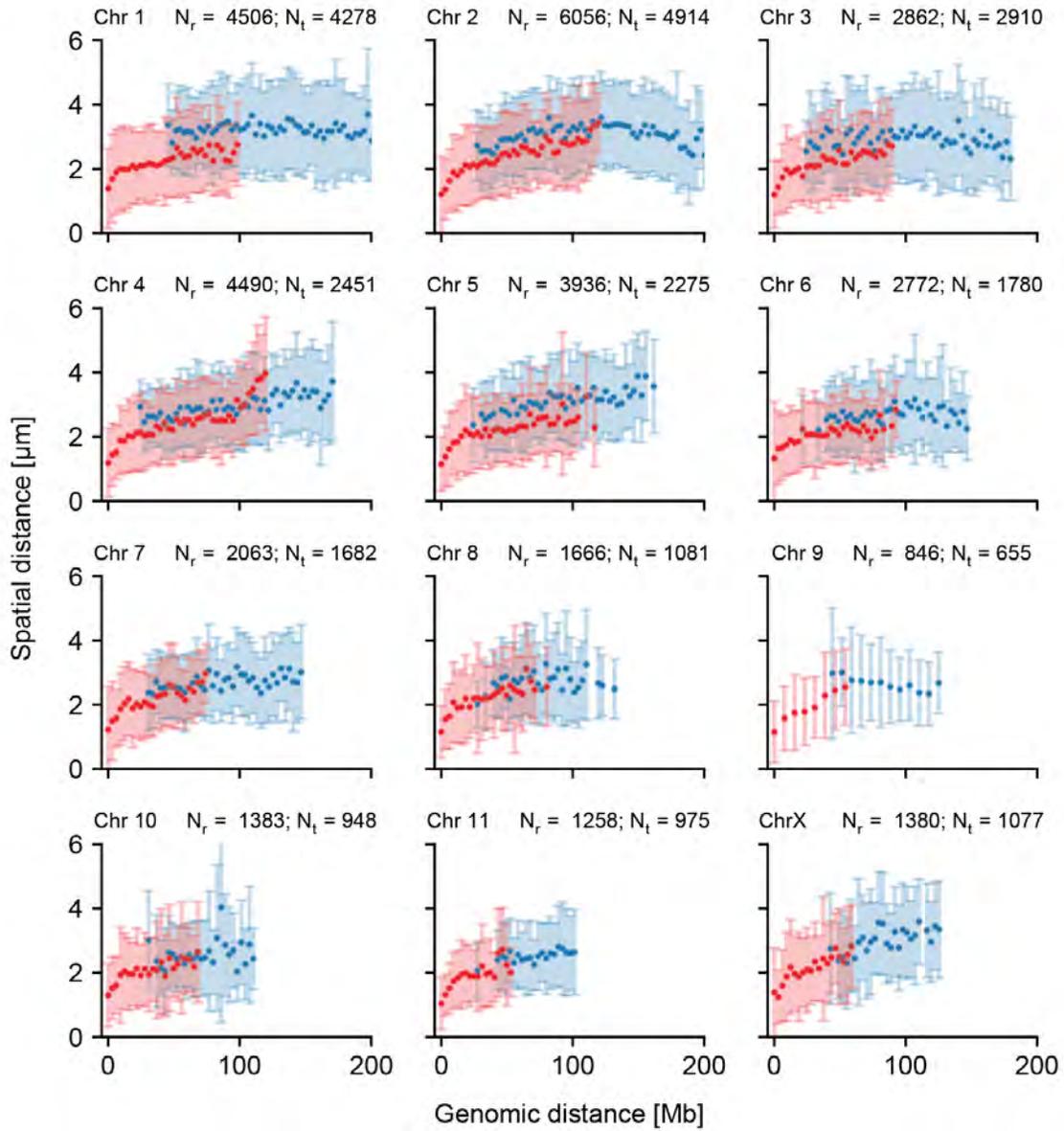
Fig. S17
Genomic vs spatial distance within and between chromosome arms, Chr 1-11 + X (PGP1f). Genomic distance versus spatial distance for Chr 1-11 & X, distinguishing between intra-arm and inter-arm measurements. Measurements were then computed, binned, and plotted as described in Fig. S15. N_r, intra-arm measurements; N_t, inter-arm measurements. The mean of each bin ± 1 SD is plotted.
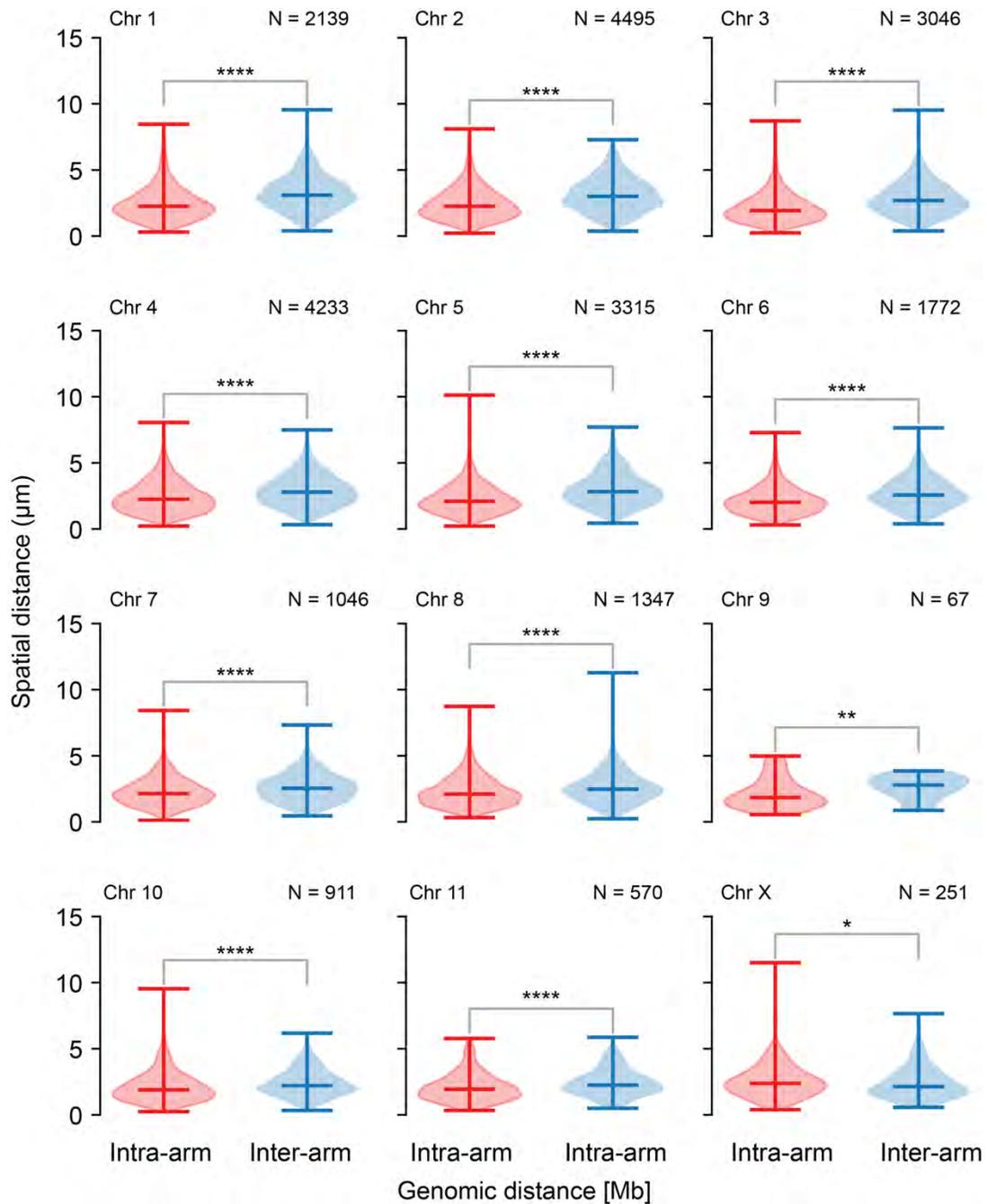
Fig. S18
Distribution of intra-arm and inter-arm pairwise distances, Chr 1-11 + X (PGP1f).
Distribution of intra-arm and inter-arm pairwise distances in genomic distance ranges shared by both types of measurement. Bins in Fig. S17 with at least 20 instances of each type of measurement were considered. There is a significant difference between the two distributions for each chromosome, with higher mean inter-arm distances in general. Chr 9 is sparse due to the idiosyncrasies described above. N, number of pairwise distance measurements in the shared genomic distance range. ****, $p < 0.0001$, ***, $p < 0.001$, **, $p < 0.01$, *, $p < 0.05$; all significance comparisons by K-S test.
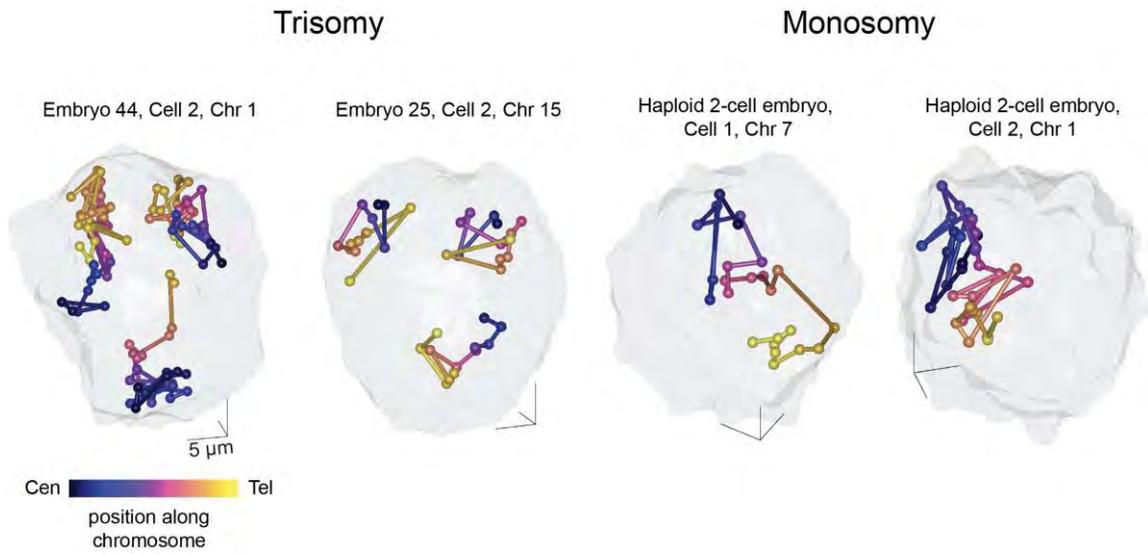
Fig. S19
Aneuploidy in embryonic cells. Examples are shown of trisomy and monosomy in four different embryonic cells, detected by IGS.
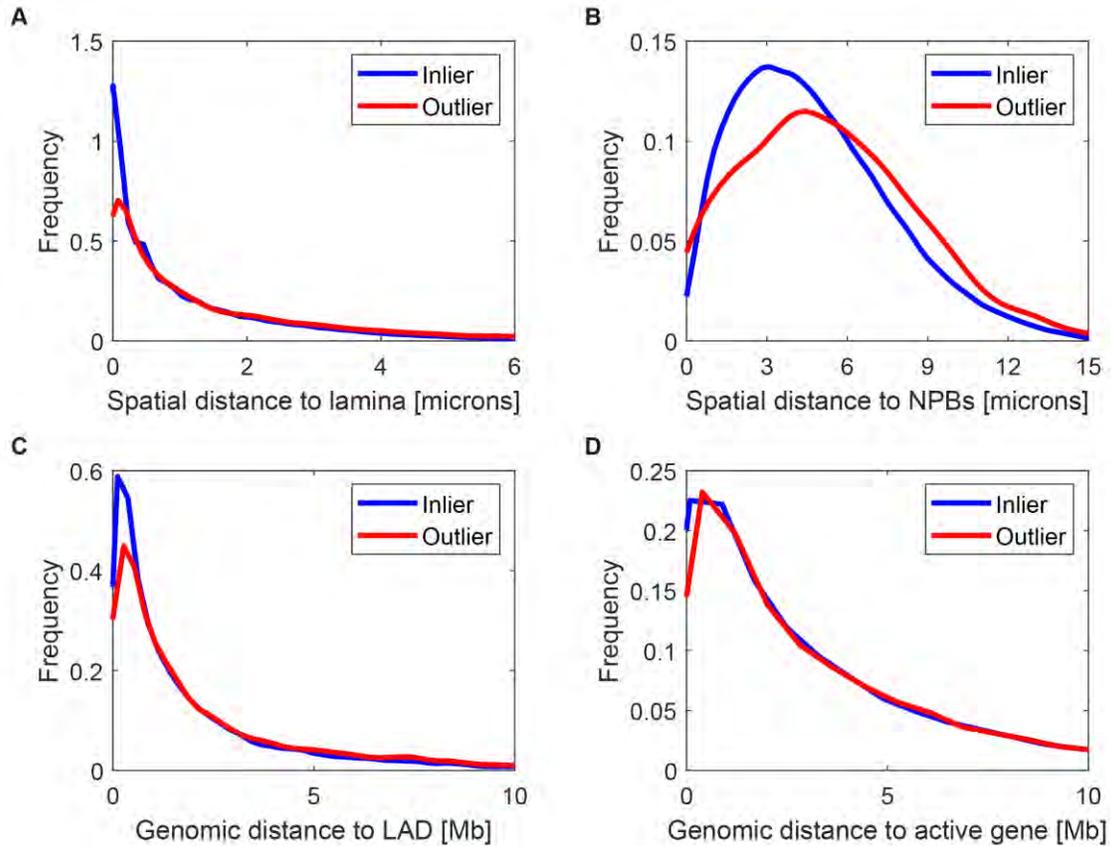
Fig. S20

Association of embryo non-colocalizing loci with nuclear landmarks and genomic annotations

(**A**) Density plot showing how inlier points (reads that colocalize with a chromosome territory) and outlier points (reads that do not colocalize) differ in their spatial proximity to the nuclear lamina. Distributions are significantly different (K-S test, $p < 10^{-51}$). (**B**) Same as (A) but for spatial distance to the nearest nucleolus precursor body (NPB). Distributions are significantly different (K-S test, $p < 10^{-109}$). (**C**) Same as (A) but for genomic distance to the nearest lamina-associated domain, identified via DamID. Distributions are significantly different (K-S test, $p < 10^{-12}$). (**D**) Same as (A) but for genomic distance to the nearest highly-expressed gene in the corresponding developmental stage. Distributions are not significantly different (K-S test, $p > 0.01$).
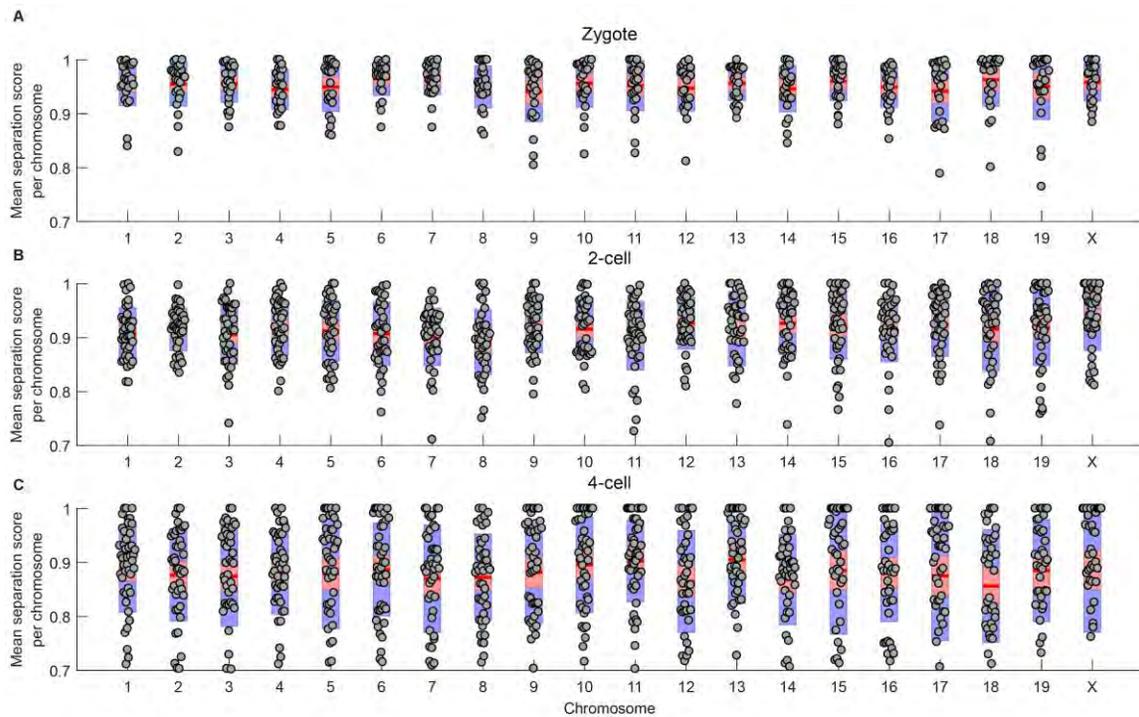
Fig. S21
Haplotype separation scores across chromosomes and developmental stages. (**A**) Boxplots showing mean (red line), 95% confidence interval (red box), and 1 SD (blue box) to represent the distribution of separation scores (**Methods**) for each chromosome in zygotes. Each overlaid grey dot represents the separation score averaged over all spatially-localized reads mapped to the associated chromosome in a single cell. (**B**) Same as (A) but for the 2-cell stage. (**C**) Same as (A) but for the 4-cell stage.
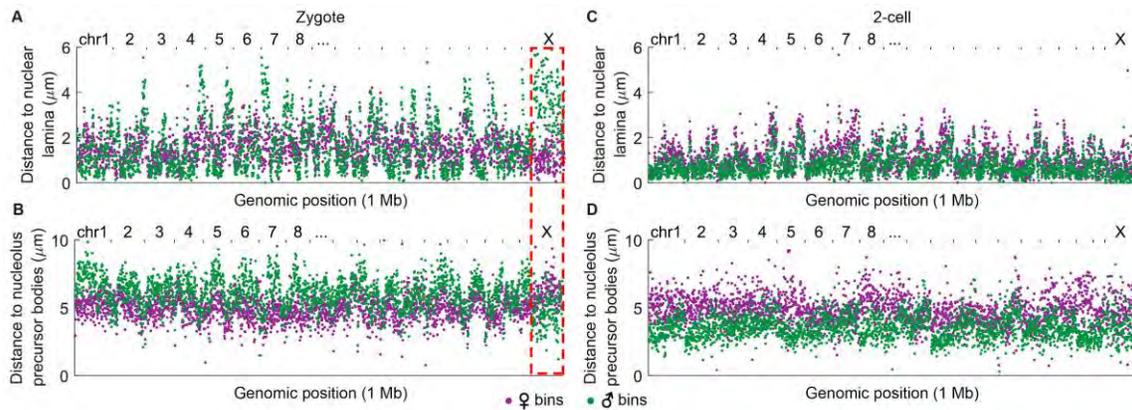
Fig. S22

Spatial distance to nuclear landmarks in the zygote and 2-cell embryos. (**A**) 1 Mb haplotype-resolved genomic bins in the zygote, plotted by position in the mm10 genome and the average spatial distance of reads in the bin to the nuclear lamina, colored by parent-of-origin. The dotted red box highlights the unique spatial localization of the paternal X chromosome away from the nuclear lamina and close to a nucleolus precursor body. (**B**) Same as (AB), but for the average distance of reads in the bin to the nearest nucleolus precursor body. (**C**) Same as (A), but for the 2-cell stage. (**D**) Same as (A), but for the average spatial distance of reads in the bin to the nearest nucleolus precursor body in the 2-cell stage.
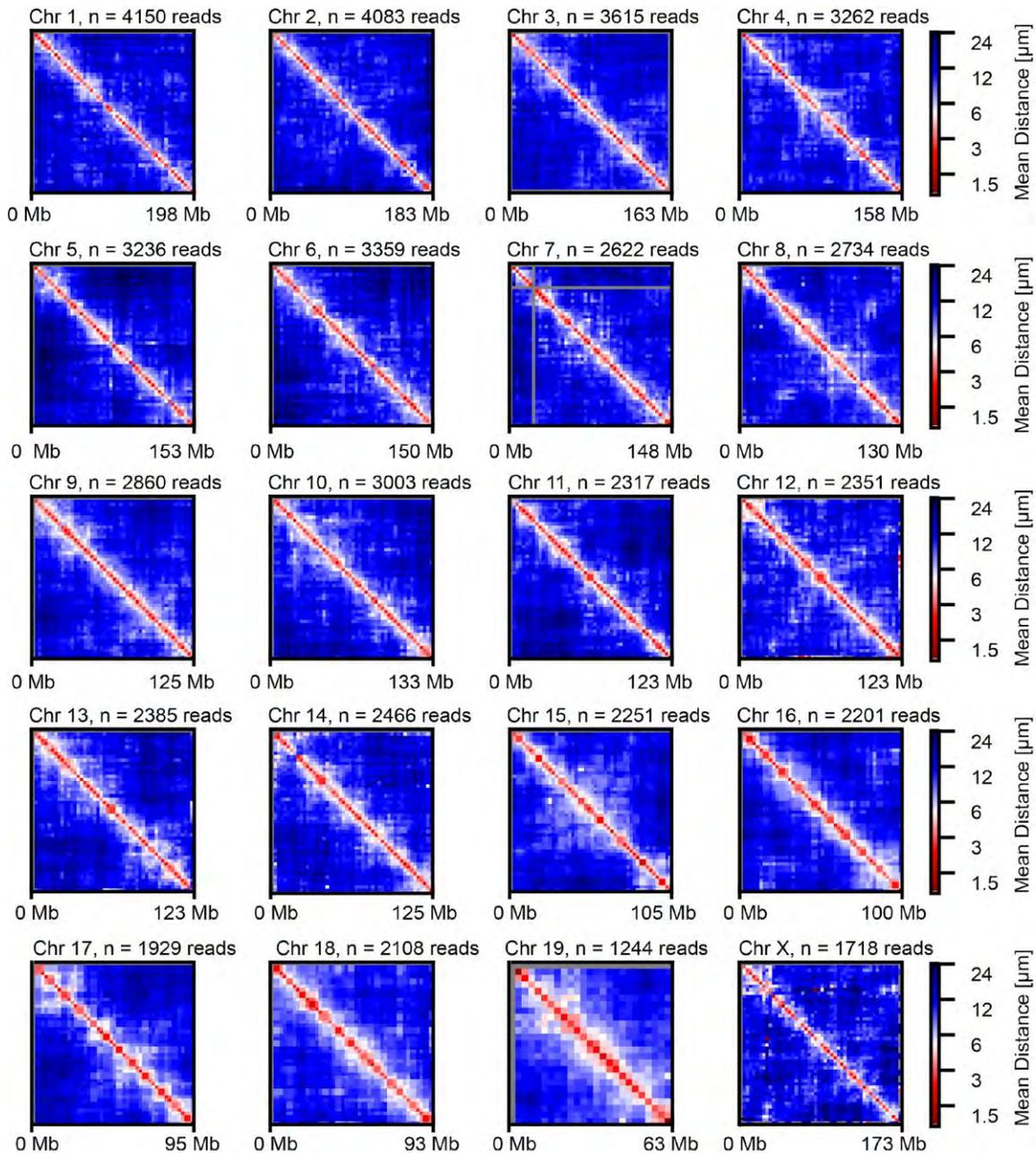
Fig. S23

Paternal zygotic ensemble distance matrices across all chromosomes. Mean spatial distance matrices for the ensemble of paternal zygotic chromosomes. In comparison to single-cell matrices (e.g. Fig. 5E), the ensemble matrices show little in the way of off-diagonal structure. Number of reads analyzed for each chromosome indicated. All autosomes describe a population ensemble of 24 pronuclei, except for ChrX, which describes an ensemble of 14 pronuclei.

44
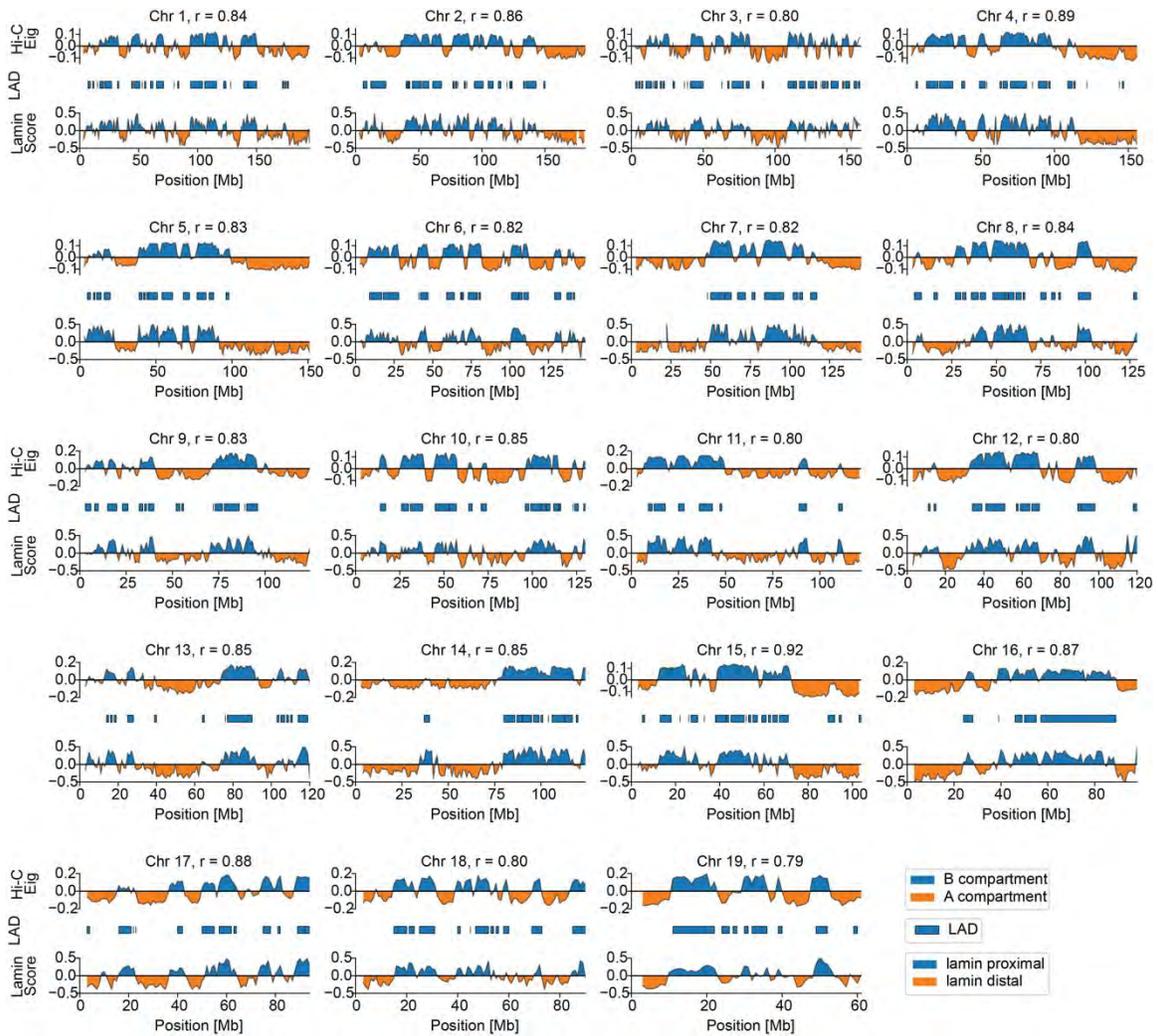
Fig. S24

Comparison of IGS with Hi-C and DamID in the paternal zygote. Each subplot shows, for a specific chromosome in paternal zygotes, Hi-C-defined compartmental status (top), DamID-defined LAD status (mid), and IGS-defined lamin proximity (bottom). Pearson correlation coefficient between lamin proximity score and Hi-C eigenvalue is indicated. Number of reads analyzed for each chromosome indicated. All analyses describe a population ensemble of 24 pronuclei. Chr X not indicated due to lack of DamID data.
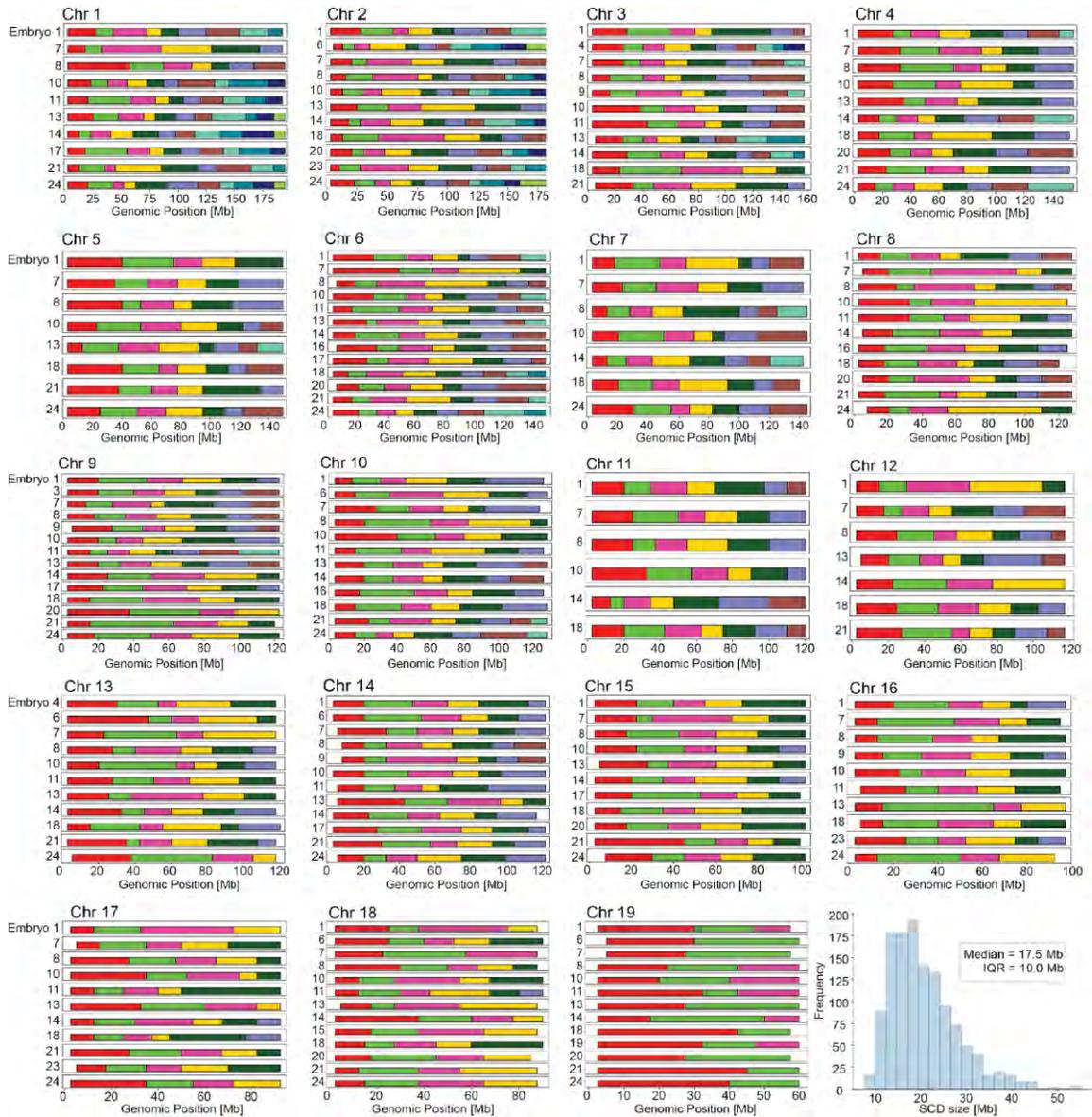
Fig. S25
Single-cell domains have heterogeneous sizes and boundaries. For all high-coverage (>90%) paternal zygotic autosomes analyzed at a matrix resolution of 2.5 Mb, SCDs were detected at that resolution (**Methods**) and are indicated. Lower right, histogram of 1262 detected SCDs spanning two or more bins. Median SCD size and IQR are indicated.

46
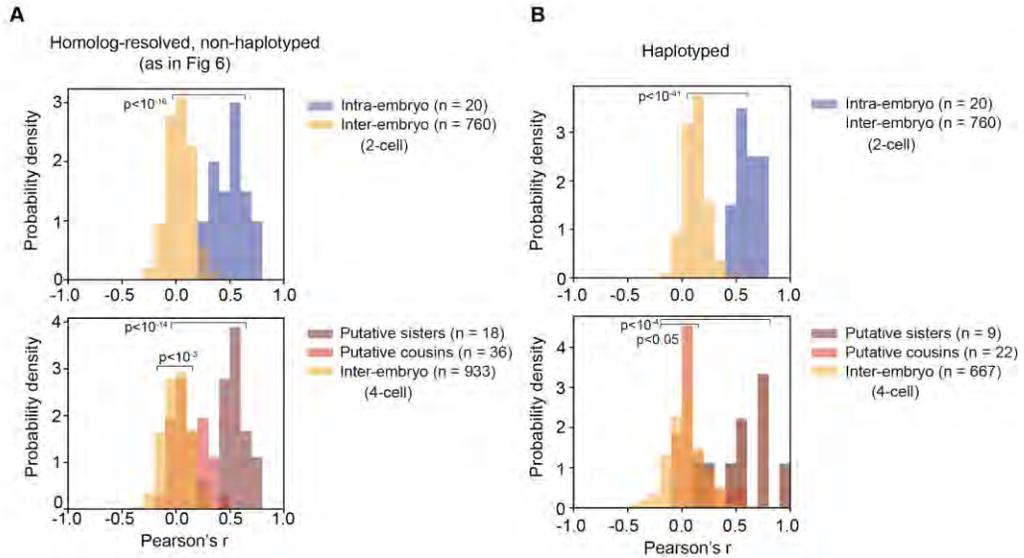
Fig. S26

Correlation of global autosome positioning for pairs of blastomeres. (**A**) Histogram representation of the data in Fig. 6C (top) and 6F (bottom). Correlations of global autosome positioning for intra- and inter-embryonic pairs of cells for 2-cell embryos (top), and for putative sister, putative cousin, and inter-embryonic pairs of cells for 4-cell embryos (bottom). Correlations were calculated in a homolog-resolved but non-haplotyped manner (**Methods**). (**B**) As in (A), but with correlations calculated in a haplotype-resolved manner (**Methods**). The number of pairs compared in each category is indicated in the legend. K-S test was used to test for significance in all cases.
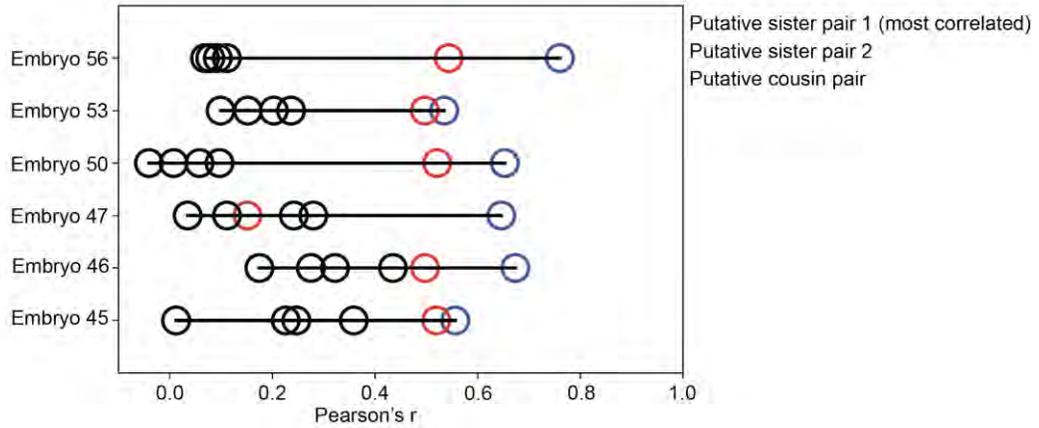
Fig. S27
Correlations of putative sister and cousin pairs for each complete four-cell embryo.
Correlations for each pair of cells in each four-cell embryo for which the genomes of all four cells had >150 reads . The putative sister pair 1 is the most correlated pair in the embryo (blue), and putative sister pair 2 is the other non-overlapping pair in the embryo (red).

Captions for Tables S1 to S4

Table S1 (provided as Excel file)
In situ sequencing dataset and annotations for PGP1f.

Table S2 (provided as Excel file)
In situ sequencing dataset and annotations for mouse early embryos.

Table S3 (provided as Excel file)
Description of data types in Tables S1 and S2.

Table S4 (provided as Excel file)
Oligonucleotides used in this study.

Table S5

Oligonucleotide sequences for SOLiD oligos. Sequences are written using the oligonucleotide modification key from Integrated DNA Technologies.

Sequence	Dye	Dibase encoded
/56-FAM//ideoxyI//ideoxyI//i3Thio-dI/NNNAA	FAM	AA
/56-FAM//ideoxyI//ideoxyI//i3Thio-dI/NNNCC	FAM	CC
/56-FAM//ideoxyI//ideoxyI//i3Thio-dI/NNNGG	FAM	GG
/56-FAM//ideoxyI//ideoxyI//i3Thio-dI/NNNTT	FAM	TT
/5Cy3/ /ideoxyI//ideoxyI//i3Thio-dI/NNNAC	Cy3	AC
/5Cy3/ /ideoxyI//ideoxyI//i3Thio-dI/NNNCA	Cy3	CA
/5Cy3/ /ideoxyI//ideoxyI//i3Thio-dI/NNNGT	Cy3	GT
/5Cy3/ /ideoxyI//ideoxyI//i3Thio-dI/NNNTG	Cy3	TG
/5TexRd-XN//ideoxyI//ideoxyI//i3Thio-dI/NNNAG	TXR	AG
/5TexRd-XN//ideoxyI//ideoxyI//i3Thio-dI/NNNGA	TXR	GA
/5TexRd-XN//ideoxyI//ideoxyI//i3Thio-dI/NNNCT	TXR	CT
/5TexRd-XN//ideoxyI//ideoxyI//i3Thio-dI/NNNTC	TXR	TC
/5Cy5/ /ideoxyI//ideoxyI//i3Thio-dI/NNNAT	Cy5	AT
/5Cy5/ /ideoxyI//ideoxyI//i3Thio-dI/NNNTA	Cy5	TA
/5Cy5/ /ideoxyI//ideoxyI//i3Thio-dI/NNNCG	Cy5	CG
/5Cy5/ /ideoxyI//ideoxyI//i3Thio-dI/NNNGC	Cy5	GC

Table S6

Replacement for SOLiD Buffer B. *This buffer excludes chloride ions, which will precipitate with the silver ions left over from the first cleavage buffer. As such, HCl should not be used for adjusting pH.

Replacement for SOLiD Buffer B	
Reagent	**Concentration**
MESNA (2-mercaptoethanolsulfate)	50 mM
tri-sodium citrate	30 mM
sodium acetate	300 mM
*pH to 7.5	

Table S7

IGS cost breakdown. Cost breakdown for the most costly reagents required to perform IGS using 300 µl total volume for each incubation. *Our Tn5 transposase was a gift from stocks produced in-house by a colleague. The estimated cost given here is an estimate for using Tn5 transposase purchased from Lucigen.

Reagent	Cost per iteration	Iterations	Cost
Tn5 transposase, (Lucigen Cat # TNP92110)*	519.60	1	519.60
Ampligase DNA ligase	3.41	1	3.41
Phusion DNA polymerase	8.56	1	8.56
EquiPhi29 DNA polymerase	65.64	1	65.64
BSPEG(9)	16.09	1	16.09
SOLiD sequencing oligo	1.31	20	26.25
T4 DNA ligase	7.86	20	157.20
Quick CIP	2.58	16	41.28
NovaSeq S2 sequencing, 2x 150 bp	5300.00	1	5300.00
Total cost			**6138.03**

Captions for Movies S1 to S5

Movie S1
Sequential rounds of in situ sequencing in PGP1f cell 85.

Movie S2
Visualization of spatially-localized reads in PGP1f cell 85.

Movie S3
Visualization of chromosome territories in PGP1f cell 85.

Movie S4
Visualization of spatially-localized reads in Embryo 21.

Movie S5
Visualization of chromosome territories in Embryo 21.

References and Notes

1. B. Bonev, G. Cavalli, Organization and function of the 3D genome. *Nat. Rev. Genet.* **17**, 661–678 (2016). doi:10.1038/nrg.2016.112 Medline

2. M. J. Rowley, V. G. Corces, Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* **19**, 789–800 (2018). doi:10.1038/s41576-018-0060-8 Medline

3. H. Zheng, W. Xie, The role of 3D genome organization in development and cell differentiation. *Nat. Rev. Mol. Cell Biol.* **20**, 535–550 (2019). doi:10.1038/s41580-019-0132-4 Medline

4. M. Spielmann, D. G. Lupiáñez, S. Mundlos, Structural variation in the 3D genome. *Nat. Rev. Genet.* **19**, 453–467 (2018). doi:10.1038/s41576-018-0007-0 Medline

5. J. Dekker, K. Rippe, M. Dekker, N. Kleckner, Capturing chromosome conformation. *Science* **295**, 1306–1311 (2002). doi:10.1126/science.1067799 Medline

6. E. Lieberman-Aiden, N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, J. Dekker, Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009). doi:10.1126/science.1181369 Medline

7. J. M. Levsky, R. H. Singer, Fluorescence in situ hybridization: Past, present and future. *J. Cell Sci.* **116**, 2833–2838 (2003). doi:10.1242/jcs.00633 Medline

8. J. R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, B. Ren, Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012). doi:10.1038/nature11082 Medline

9. T. Sexton, E. Yaffe, E. Kenigsberg, F. Bantignies, B. Leblanc, M. Hoichman, H. Parrinello, A. Tanay, G. Cavalli, Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**, 458–472 (2012). doi:10.1016/j.cell.2012.01.010 Medline

10. S. S. P. Rao, M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, E. L. Aiden, A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014). doi:10.1016/j.cell.2014.11.021 Medline

11. E. P. Nora, B. R. Lajoie, E. G. Schulz, L. Giorgetti, I. Okamoto, N. Servant, T. Piolot, N. L. van Berkum, J. Meisig, J. Sedat, J. Gribnau, E. Barillot, N. Blüthgen, J. Dekker, E. Heard, Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012). doi:10.1038/nature11049 Medline

12. R. Kalhor, H. Tjong, N. Jayathilaka, F. Alber, L. Chen, Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. *Nat. Biotechnol.* **30**, 90–98 (2011). doi:10.1038/nbt.2057 Medline

13. T. Nagano, Y. Lubling, T. J. Stevens, S. Schoenfelder, E. Yaffe, W. Dean, E. D. Laue, A. Tanay, P. Fraser, Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502**, 59–64 (2013). doi:10.1038/nature12593 Medline

14. T. Nagano, Y. Lubling, C. Várnai, C. Dudley, W. Leung, Y. Baran, N. Mendelson Cohen, S.

Wingett, P. Fraser, A. Tanay, Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature* **547**, 61–67 (2017). doi:10.1038/nature23001 Medline

15. V. Ramani, X. Deng, R. Qiu, K. L. Gunderson, F. J. Steemers, C. M. Disteche, W. S. Noble, Z. Duan, J. Shendure, Massively multiplex single-cell Hi-C. *Nat. Methods* **14**, 263–266 (2017). doi:10.1038/nmeth.4155 Medline

16. T. J. Stevens, D. Lando, S. Basu, L. P. Atkinson, Y. Cao, S. F. Lee, M. Leeb, K. J. Wohlfahrt, W. Boucher, A. O'Shaughnessy-Kirwan, J. Cramard, A. J. Faure, M. Ralser, E. Blanco, L. Morey, M. Sansó, M. G. S. Palayret, B. Lehner, L. Di Croce, A. Wutz, B. Hendrich, D. Klenerman, E. D. Laue, 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* **544**, 59–64 (2017). doi:10.1038/nature21429 Medline

17. L. Tan, D. Xing, C.-H. Chang, H. Li, X. S. Xie, Three-dimensional genome structures of single diploid human cells. *Science* **361**, 924–928 (2018). doi:10.1126/science.aat5641 Medline

18. T. Cremer, M. Cremer, Chromosome territories. *Cold Spring Harb. Perspect. Biol.* **2**, a003889 (2010). doi:10.1101/cshperspect.a003889 Medline

19. J. Fraser, I. Williamson, W. A. Bickmore, J. Dostie, An Overview of Genome Organization and How We Got There: From FISH to Hi-C. *Microbiol. Mol. Biol. Rev.* **79**, 347–372 (2015). doi:10.1128/MMBR.00006-15 Medline

20. S. Wang, J.-H. Su, B. J. Beliveau, B. Bintu, J. R. Moffitt, C. T. Wu, X. Zhuang, Spatial organization of chromatin domains and compartments in single chromosomes. *Science* **353**, 598–602 (2016). doi:10.1126/science.aaf8084 Medline

21. D. I. Cattoni, A. M. Cardozo Gizzi, M. Georgieva, M. Di Stefano, A. Valeri, D. Chamousset, C. Houbron, S. Déjardin, J.-B. Fiche, I. González, J.-M. Chang, T. Sexton, M. A. Marti-Renom, F. Bantignies, G. Cavalli, M. Nollmann, Single-cell absolute contact probability detection reveals chromosomes are organized by multiple low-frequency yet specific interactions. *Nat. Commun.* **8**, 1753 (2017). doi:10.1038/s41467-017-01962-x Medline

22. B. Bintu, L. J. Mateo, J.-H. Su, N. A. Sinnott-Armstrong, M. Parker, S. Kinrot, K. Yamaya, A. N. Boettiger, X. Zhuang, Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science* **362**, eaau1783 (2018). doi:10.1126/science.aau1783 Medline

23. G. Nir, I. Farabella, C. Pérez Estrada, C. G. Ebeling, B. J. Beliveau, H. M. Sasaki, S. D. Lee, S. C. Nguyen, R. B. McCole, S. Chattoraj, J. Erceg, J. AlHaj Abed, N. M. C. Martins, H. Q. Nguyen, M. A. Hannan, S. Russell, N. C. Durand, S. S. P. Rao, J. Y. Kishi, P. Soler-Vila, M. Di Pierro, J. N. Onuchic, S. P. Callahan, J. M. Schreiner, J. A. Stuckey, P. Yin, E. L. Aiden, M. A. Marti-Renom, C. T. Wu, Walking along chromosomes with super-resolution imaging, contact maps, and integrative modeling. *PLOS Genet.* **14**, e1007872 (2018). doi:10.1371/journal.pgen.1007872 Medline

24. Q. Szabo, D. Jost, J.-M. Chang, D. I. Cattoni, G. L. Papadopoulos, B. Bonev, T. Sexton, J. Gurgo, C. Jacquier, M. Nollmann, F. Bantignies, G. Cavalli, TADs are 3D structural units of higher-order chromosome organization in *Drosophila*. *Sci. Adv.* **4**, eaar8082 (2018). doi:10.1126/sciadv.aar8082 Medline

25. E. H. Finn, G. Pegoraro, H. B. Brandão, A.-L. Valton, M. E. Oomen, J. Dekker, L. Mirny, T.

Misteli; Extensive Heterogeneity and Intrinsic Variation in Spatial Genome Organization, Extensive heterogeneity and intrinsic variation in spatial genome organization. *Cell* **176**, 1502–1515.e10 (2019). doi:10.1016/j.cell.2019.01.020 Medline

26. L. J. Mateo, S. E. Murphy, A. Hafner, I. S. Cinquini, C. A. Walker, A. N. Boettiger, Visualizing DNA folding and RNA in embryos at single-cell resolution. *Nature* **568**, 49–54 (2019). doi:10.1038/s41586-019-1035-4 Medline

27. A. M. Cardozo Gizzi, D. I. Cattoni, J.-B. Fiche, S. M. Espinola, J. Gurgo, O. Messina, C. Houbron, Y. Ogiyama, G. L. Papadopoulos, G. Cavalli, M. Lagha, M. Nollmann, Microscopy-based chromosome conformation capture enables simultaneous visualization of genome organization and transcription in intact organisms. *Mol. Cell* **74**, 212–222.e5 (2019). doi:10.1016/j.molcel.2019.01.011 Medline

28. B. J. Beliveau, A. N. Boettiger, M. S. Avendaño, R. Jungmann, R. B. McCole, E. F. Joyce, C. Kim-Kiselak, F. Bantignies, C. Y. Fonseka, J. Erceg, M. A. Hannan, H. G. Hoang, D. Colognori, J. T. Lee, W. M. Shih, P. Yin, X. Zhuang, C. T. Wu, Single-molecule super-resolution imaging of chromosomes and in situ haplotype visualization using Oligopaint FISH probes. *Nat. Commun.* **6**, 7147 (2015). doi:10.1038/ncomms8147 Medline

29. Materials and methods are available as supplementary materials.

30. X. Chen, Y. Shen, W. Draper, J. D. Buenrostro, U. Litzenburger, S. W. Cho, A. T. Satpathy, A. C. Carter, R. P. Ghosh, A. East-Seletsky, J. A. Doudna, W. J. Greenleaf, J. T. Liphardt, H. Y. Chang, ATAC-see reveals the accessible genome by transposase-mediated imaging and sequencing. *Nat. Methods* **13**, 1013–1020 (2016). doi:10.1038/nmeth.4031 Medline

31. J. H. Lee, E. R. Daugharthy, J. Scheiman, R. Kalhor, J. L. Yang, T. C. Ferrante, R. Terry, S. S. F. Jeanty, C. Li, R. Amamoto, D. T. Peters, B. M. Turczyk, A. H. Marblestone, S. A. Inverso, A. Bernard, P. Mali, X. Rios, J. Aach, G. M. Church, Highly multiplexed subcellular RNA sequencing in situ. *Science* **343**, 1360–1363 (2014). doi:10.1126/science.1250212 Medline

32. J. R. Moffitt, J. Hao, G. Wang, K. H. Chen, H. P. Babcock, X. Zhuang, High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 11046–11051 (2016). doi:10.1073/pnas.1612826113 Medline

33. E. V. Volpi, E. Chevret, T. Jones, R. Vatcheva, J. Williamson, S. Beck, R. D. Campbell, M. Goldsworthy, S. H. Powis, J. Ragoussis, J. Trowsdale, D. Sheer, Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. *J. Cell Sci.* **113**, 1565–1576 (2000). Medline

34. A. Bolzer, G. Kreth, I. Solovei, D. Koehler, K. Saracoglu, C. Fauth, S. Müller, R. Eils, C. Cremer, M. R. Speicher, T. Cremer, Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLOS Biol.* **3**, e157 (2005). doi:10.1371/journal.pbio.0030157 Medline

35. A. P. J. de Koning, W. Gu, T. A. Castoe, M. A. Batzer, D. D. Pollock, Repetitive elements may comprise over two-thirds of the human genome. *PLOS Genet.* **7**, e1002384 (2011). doi:10.1371/journal.pgen.1002384 Medline

36. K. H. Miga, S. Koren, A. Rhie, M. R. Vollger, A. Gershman, A. Bzikadze, S. Brooks, E. Howe, D. Porubsky, G. A. Logsdon, V. A. Schneider, T. Potapova, J. Wood, W. Chow, J. Armstrong, J. Fredrickson, E. Pak, K. Tigyi, M. Kremitzki, C. Markovic, V. Maduro, A. Dutra, G. G. Bouffard, A. M. Chang, N. F. Hansen, A. B. Wilfert, F. Thibaud-Nissen, A. D. Schmitt, J.-M. Belton, S. Selvaraj, M. Y. Dennis, D. C. Soto, R. Sahasrabudhe, G. Kaya, J. Quick, N. J. Loman, N. Holmes, M. Loose, U. Surti, R. A. Risques, T. A. Graves Lindsay, R. Fulton, I. Hall, B. Paten, K. Howe, W. Timp, A. Young, J. C. Mullikin, P. A. Pevzner, J. L. Gerton, B. A. Sullivan, E. E. Eichler, A. M. Phillippy, Telomere-to-telomere assembly of a complete human X chromosome. *Nature* **585**, 79–84 (2020). doi:10.1038/s41586-020-2547-7 Medline

37. A. Cournac, R. Koszul, J. Mozziconacci, The 3D folding of metazoan genomes correlates with the association of similar repetitive elements. *Nucleic Acids Res.* **44**, 245–255 (2016). doi:10.1093/nar/gkv1292 Medline

38. V. Casa, D. Gabellini, A repetitive elements perspective in Polycomb epigenetics. *Front. Genet.* **3**, 199 (2012). doi:10.3389/fgene.2012.00199 Medline

39. M. Hausmann, J.-H. Lee, A. Sievers, M. Krufczik, G. Hildenbrand, COMBinatorial Oligonucleotide FISH (COMBO-FISH) with uniquely binding repetitive DNA probes. *Methods Mol. Biol.* **2175**, 65–77 (2020). doi:10.1007/978-1-0716-0763-3_6 Medline

40. W. Bao, K. K. Kojima, O. Kohany, Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob. DNA* **6**, 11 (2015). doi:10.1186/s13100-015-0041-9 Medline

41. M. Falk, Y. Feodorova, N. Naumova, M. Imakaev, B. R. Lajoie, H. Leonhardt, B. Joffe, J. Dekker, G. Fudenberg, I. Solovei, L. A. Mirny, Heterochromatin drives compartmentalization of inverted and conventional nuclei. *Nature* **570**, 395–399 (2019). doi:10.1038/s41586-019-1275-3 Medline

42. H. Muller, J. Gil Jr., I. A. Drinnenberg, The impact of centromeres on spatial genome architecture. *Trends Genet.* **35**, 565–578 (2019). doi:10.1016/j.tig.2019.05.003 Medline

43. C. B. Hug, J. M. Vaquerizas, The birth of the 3D genome during early embryonic development. *Trends Genet.* **34**, 903–914 (2018). doi:10.1016/j.tig.2018.09.002 Medline

44. C. Chazaud, Y. Yamanaka, Lineage specification in the mouse preimplantation embryo. *Development* **143**, 1063–1074 (2016). doi:10.1242/dev.128314 Medline

45. S. H. Namekawa, B. Payer, K. D. Huynh, R. Jaenisch, J. T. Lee, Two-step imprinted X inactivation: Repeat versus genic silencing in the mouse. *Mol. Cell. Biol.* **30**, 3187–3205 (2010). doi:10.1128/MCB.00227-10 Medline

46. M. Borsos, S. M. Perricone, T. Schauer, J. Pontabry, K. L. de Luca, S. S. de Vries, E. R. Ruiz-Morales, M.-E. Torres-Padilla, J. Kind, Genome-lamina interactions are established de novo in the early mouse embryo. *Nature* **569**, 729–733 (2019). doi:10.1038/s41586-019-1233-0 Medline

47. I. M. Flyamer, J. Gassler, M. Imakaev, H. B. Brandão, S. V. Ulianov, N. Abdennur, S. V. Razin, L. A. Mirny, K. Tachibana-Konwalski, Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature* **544**, 110–114 (2017). doi:10.1038/nature21711 Medline

48. Z. Du, H. Zheng, B. Huang, R. Ma, J. Wu, X. Zhang, J. He, Y. Xiang, Q. Wang, Y. Li, J. Ma,

X. Zhang, K. Zhang, Y. Wang, M. Q. Zhang, J. Gao, J. R. Dixon, X. Wang, J. Zeng, W. Xie, Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* **547**, 232–235 (2017). doi:10.1038/nature23263 Medline

49. Y. Ke, Y. Xu, X. Chen, S. Feng, Z. Liu, Y. Sun, X. Yao, F. Li, W. Zhu, L. Gao, H. Chen, Z. Du, W. Xie, X. Xu, X. Huang, J. Liu, 3D chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis. *Cell* **170**, 367–381.e20 (2017). doi:10.1016/j.cell.2017.06.029 Medline

50. S. Collombet, N. Ranisavljevic, T. Nagano, C. Varnai, T. Shisode, W. Leung, T. Piolot, R. Galupa, M. Borensztein, N. Servant, P. Fraser, K. Ancelin, E. Heard, Parental-to-embryo switch of chromosome organization in early embryogenesis. *Nature* **580**, 142–146 (2020). doi:10.1038/s41586-020-2125-z Medline

51. M.-E. Torres-Padilla, D.-E. Parfitt, T. Kouzarides, M. Zernicka-Goetz, Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature* **445**, 214–218 (2007). doi:10.1038/nature05458 Medline

52. A. Burton, J. Muller, S. Tu, P. Padilla-Longoria, E. Guccione, M.-E. Torres-Padilla, Single-cell profiling of epigenetic modifiers identifies PRDM14 as an inducer of cell fate in the mammalian embryo. *Cell Rep.* **5**, 687–701 (2013). doi:10.1016/j.celrep.2013.09.044 Medline

53. A. Burton, M.-E. Torres-Padilla, Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. *Nat. Rev. Mol. Cell Biol.* **15**, 723–735 (2014). doi:10.1038/nrm3885 Medline

54. T. Aguirre-Lavin, P. Adenot, A. Bonnet-Garnier, G. Lehmann, R. Fleurot, C. Boulesteix, P. Debey, N. Beaujean, 3D-FISH analysis of embryonic nuclei in mouse highlights several abrupt changes of nuclear organization during preimplantation development. *BMC Dev. Biol.* **12**, 30 (2012). doi:10.1186/1471-213X-12-30 Medline

55. H. Fulka, F. Aoki, Nucleolus precursor bodies and ribosome biogenesis in early mammalian embryos: Old theories and new discoveries. *Biol. Reprod.* **94**, 143 (2016). doi:10.1095/biolreprod.115.136093 Medline

56. W. Mayer, A. Smith, R. Fundele, T. Haaf, Spatial separation of parental genomes in preimplantation mouse embryos. *J. Cell Biol.* **148**, 629–634 (2000). doi:10.1083/jcb.148.4.629 Medline

57. J. Reichmann, B. Nijmeijer, M. J. Hossain, M. Eguren, I. Schneider, A. Z. Politi, M. J. Roberti, L. Hufnagel, T. Hiiragi, J. Ellenberg, Dual-spindle formation in zygotes keeps parental genomes apart in early mammalian embryos. *Science* **361**, 189–193 (2018). doi:10.1126/science.aar7462 Medline

58. C. R. Cowan, P. M. Carlton, W. Z. Cande; Rabl Organization and the Bouquet, The polar arrangement of telomeres in interphase and meiosis. Rabl organization and the bouquet. *Plant Physiol.* **125**, 532–538 (2001). doi:10.1104/pp.125.2.532 Medline

59. M. Pouokam, B. Cruz, S. Burgess, M. R. Segal, M. Vazquez, J. Arsuaga, The Rabl configuration limits topological entanglement of chromosomes in budding yeast. *Sci. Rep.* **9**, 6795 (2019). doi:10.1038/s41598-019-42967-4 Medline

60. M. T. Lee, A. R. Bonneau, A. J. Giraldez, Zygotic genome activation during the maternal-to-

X. Zhang, K. Zhang, Y. Wang, M. Q. Zhang, J. Gao, J. R. Dixon, X. Wang, J. Zeng, W. Xie, Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* **547**, 232–235 (2017). doi:10.1038/nature23263 Medline

49. Y. Ke, Y. Xu, X. Chen, S. Feng, Z. Liu, Y. Sun, X. Yao, F. Li, W. Zhu, L. Gao, H. Chen, Z. Du, W. Xie, X. Xu, X. Huang, J. Liu, 3D chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis. *Cell* **170**, 367–381.e20 (2017). doi:10.1016/j.cell.2017.06.029 Medline

50. S. Collombet, N. Ranisavljevic, T. Nagano, C. Varnai, T. Shisode, W. Leung, T. Piolot, R. Galupa, M. Borensztein, N. Servant, P. Fraser, K. Ancelin, E. Heard, Parental-to-embryo switch of chromosome organization in early embryogenesis. *Nature* **580**, 142–146 (2020). doi:10.1038/s41586-020-2125-z Medline

51. M.-E. Torres-Padilla, D.-E. Parfitt, T. Kouzarides, M. Zernicka-Goetz, Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature* **445**, 214–218 (2007). doi:10.1038/nature05458 Medline

52. A. Burton, J. Muller, S. Tu, P. Padilla-Longoria, E. Guccione, M.-E. Torres-Padilla, Single-cell profiling of epigenetic modifiers identifies PRDM14 as an inducer of cell fate in the mammalian embryo. *Cell Rep.* **5**, 687–701 (2013). doi:10.1016/j.celrep.2013.09.044 Medline

53. A. Burton, M.-E. Torres-Padilla, Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. *Nat. Rev. Mol. Cell Biol.* **15**, 723–735 (2014). doi:10.1038/nrm3885 Medline

54. T. Aguirre-Lavin, P. Adenot, A. Bonnet-Garnier, G. Lehmann, R. Fleurot, C. Boulesteix, P. Debey, N. Beaujean, 3D-FISH analysis of embryonic nuclei in mouse highlights several abrupt changes of nuclear organization during preimplantation development. *BMC Dev. Biol.* **12**, 30 (2012). doi:10.1186/1471-213X-12-30 Medline

55. H. Fulka, F. Aoki, Nucleolus precursor bodies and ribosome biogenesis in early mammalian embryos: Old theories and new discoveries. *Biol. Reprod.* **94**, 143 (2016). doi:10.1095/biolreprod.115.136093 Medline

56. W. Mayer, A. Smith, R. Fundele, T. Haaf, Spatial separation of parental genomes in preimplantation mouse embryos. *J. Cell Biol.* **148**, 629–634 (2000). doi:10.1083/jcb.148.4.629 Medline

57. J. Reichmann, B. Nijmeijer, M. J. Hossain, M. Eguren, I. Schneider, A. Z. Politi, M. J. Roberti, L. Hufnagel, T. Hiiragi, J. Ellenberg, Dual-spindle formation in zygotes keeps parental genomes apart in early mammalian embryos. *Science* **361**, 189–193 (2018). doi:10.1126/science.aar7462 Medline

58. C. R. Cowan, P. M. Carlton, W. Z. Cande; Rabl Organization and the Bouquet, The polar arrangement of telomeres in interphase and meiosis. Rabl organization and the bouquet. *Plant Physiol.* **125**, 532–538 (2001). doi:10.1104/pp.125.2.532 Medline

59. M. Pouokam, B. Cruz, S. Burgess, M. R. Segal, M. Vazquez, J. Arsuaga, The Rabl configuration limits topological entanglement of chromosomes in budding yeast. *Sci. Rep.* **9**, 6795 (2019). doi:10.1038/s41598-019-42967-4 Medline

60. M. T. Lee, A. R. Bonneau, A. J. Giraldez, Zygotic genome activation during the maternal-to-

zygotic transition. *Annu. Rev. Cell Dev. Biol.* **30**, 581–613 (2014). doi:10.1146/annurev-cellbio-100913-013027 Medline

61. K. Ahmed, H. Dehghani, P. Rugg-Gunn, E. Fussner, J. Rossant, D. P. Bazett-Jones, Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo. *PLOS ONE* **5**, e10531 (2010). doi:10.1371/journal.pone.0010531 Medline

62. N. Plachta, T. Bollenbach, S. Pease, S. E. Fraser, P. Pantazis, Oct4 kinetics predict cell lineage patterning in the early mammalian embryo. *Nat. Cell Biol.* **13**, 117–123 (2011). doi:10.1038/ncb2154 Medline

63. K. Piotrowska-Nitsche, A. Perea-Gomez, S. Haraguchi, M. Zernicka-Goetz, Four-cell stage mouse blastomeres have different developmental properties. *Development* **132**, 479–490 (2005). doi:10.1242/dev.01602 Medline

64. J. Walter, L. Schermelleh, M. Cremer, S. Tashiro, T. Cremer, Chromosome order in HeLa cells changes during mitosis and early G1, but is stably maintained during subsequent interphase stages. *J. Cell Biol.* **160**, 685–697 (2003). doi:10.1083/jcb.200211103 Medline

65. D. Gerlich, J. Beaudouin, B. Kalbfuss, N. Daigle, R. Eils, J. Ellenberg, Global chromosome positions are transmitted through mitosis in mammalian cells. *Cell* **112**, 751–764 (2003). doi:10.1016/S0092-8674(03)00189-2 Medline

66. J. Kind, L. Pagie, H. Ortabozkoyun, S. Boyle, S. S. de Vries, H. Janssen, M. Amendola, L. D. Nolen, W. A. Bickmore, B. van Steensel, Single-cell dynamics of genome-nuclear lamina interactions. *Cell* **153**, 178–192 (2013). doi:10.1016/j.cell.2013.02.028 Medline

67. I. Thomson, S. Gilchrist, W. A. Bickmore, J. R. Chubb, The radial positioning of chromatin is not inherited through mitosis but is established de novo in early G1. *Curr. Biol.* **14**, 166–172 (2004). doi:10.1016/j.cub.2003.12.024 Medline

68. C.-S. Lee, R. W. Wang, H.-H. Chang, D. Capurso, M. R. Segal, J. E. Haber, Chromosome position determines the success of double-strand break repair. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E146–E154 (2016). doi:10.1073/pnas.1523660113 Medline

69. E. R. Phillips, P. J. McKinnon, DNA double-strand break repair and development. *Oncogene* **26**, 7799–7808 (2007). doi:10.1038/sj.onc.1210877 Medline

70. E. Vanneste, T. Voet, C. Le Caignec, M. Ampe, P. Konings, C. Melotte, S. Debrock, M. Amyere, M. Vikkula, F. Schuit, J.-P. Fryns, G. Verbeke, T. D'Hooghe, Y. Moreau, J. R. Vermeesch, Chromosome instability is common in human cleavage-stage embryos. *Nat. Med.* **15**, 577–583 (2009). doi:10.1038/nm.1924 Medline

71. Y. Zhang, R. P. McCord, Y.-J. Ho, B. R. Lajoie, D. G. Hildebrand, A. C. Simon, M. S. Becker, F. W. Alt, J. Dekker, Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell* **148**, 908–921 (2012). doi:10.1016/j.cell.2012.02.002 Medline

72. J. M. Engreitz, V. Agarwala, L. A. Mirny, Three-dimensional genome architecture influences partner selection for chromosomal translocations in human disease. *PLOS ONE* **7**, e44196 (2012). doi:10.1371/journal.pone.0044196 Medline

73. H. Q. Nguyen, S. Chattoraj, D. Castillo, S. C. Nguyen, G. Nir, A. Lioutas, E. A. Hershberg, N. M. C. Martins, P. L. Reginato, M. Hannan, B. J. Beliveau, G. M. Church, E. R.

Daugharthy, M. A. Marti-Renom, C. T. Wu, 3D mapping and accelerated super-resolution imaging of the human genome using in situ sequencing. *Nat. Methods* **17**, 822–832 (2020). doi:10.1038/s41592-020-0890-0 Medline

74. J.-H. Su, P. Zheng, S. S. Kinrot, B. Bintu, X. Zhuang, Genome-scale imaging of the 3D organization and transcriptional activity of chromatin. *Cell* **182**, 1641–1659.e26 (2020). doi:10.1016/j.cell.2020.07.032 Medline

75. J. M. Luppino, D. S. Park, S. C. Nguyen, Y. Lan, Z. Xu, R. Yunker, E. F. Joyce, Cohesin promotes stochastic domain intermingling to ensure proper regulation of boundary-proximal genes. *Nat. Genet.* **52**, 840–848 (2020). doi:10.1038/s41588-020-0647-9 Medline

76. B. D. Fields, S. C. Nguyen, G. Nir, S. Kennedy, A multiplexed DNA FISH strategy for assessing genome architecture in *Caenorhabditis elegans*. *eLife* **8**, e42823 (2019). doi:10.7554/eLife.42823 Medline

77. R. A. Beagrie, A. Scialdone, M. Schueler, D. C. A. Kraemer, M. Chotalia, S. Q. Xie, M. Barbieri, I. de Santiago, L.-M. Lavitas, M. R. Branco, J. Fraser, J. Dostie, L. Game, N. Dillon, P. A. W. Edwards, M. Nicodemi, A. Pombo, Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* **543**, 519–524 (2017). doi:10.1038/nature21411 Medline

78. S. A. Quinodoz, N. Ollikainen, B. Tabak, A. Palla, J. M. Schmidt, E. Detmar, M. M. Lai, A. A. Shishkin, P. Bhat, Y. Takei, V. Trinh, E. Aznauryan, P. Russell, C. Cheng, M. Jovanovic, A. Chow, L. Cai, P. McDonel, M. Garber, M. Guttman, Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* **174**, 744–757.e24 (2018). doi:10.1016/j.cell.2018.05.024 Medline

79. G. Girelli, J. Custodio, T. Kallas, F. Agostini, E. Wernersson, B. Spanjaard, A. Mota, S. Kolbeinsdottir, E. Gelali, N. Crosetto, M. Bienko, GPSeq reveals the radial organization of chromatin in the cell nucleus. *Nat. Biotechnol.* **38**, 1184–1193 (2020). Medline

80. D. Zink, A. H. Fischer, J. A. Nickerson, Nuclear structure in cancer cells. *Nat. Rev. Cancer* **4**, 677–687 (2004). doi:10.1038/nrc1430 Medline

81. J. D. Buenrostro, P. G. Giresi, L. C. Zaba, H. Y. Chang, W. J. Greenleaf, Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013). doi:10.1038/nmeth.2688 Medline

82. C.-M. Clausson, L. Arngården, O. Ishaq, A. Klaesson, M. Kühnemund, K. Grannas, B. Koos, X. Qian, P. Ranefall, T. Krzywkowski, H. Brismar, M. Nilsson, C. Wählby, O. Söderberg, Compaction of rolling circle amplification products increases signal integrity and signal-to-noise ratio. *Sci. Rep.* **5**, 12317 (2015). doi:10.1038/srep12317 Medline

83. F. Chen, P. W. Tillberg, E. S. Boyden, Optical imaging. Expansion microscopy. *Science* **347**, 543–548 (2015). doi:10.1126/science.1260088 Medline

84. S. Alon, D. R. Goodwin, A. Sinha, A. T. Wassie, F. Chen, E. R. Daugharthy, Y. Bando, A. Kajita, A. G. Xue, K. Marrett, R. Prior, Y. Cui, A. C. Payne, C.-C. Yao, H.-J. Suk, R. Wang, C.-C. Yu, P. Tillberg, P. Reginato, N. Pak, S. Liu, S. Punthambaker, E. P. R. Iyer, R. E. Kohman, J. A. Miller, E. S. Lein, A. Lako, N. Cullen, S. Rodig, K. Helvie, D. L.

Abravanel, N. Wagle, B. E. Johnson, J. Klughammer, M. Slyper, J. Waldman, J. Jané-Valbuena, O. Rozenblatt-Rosen, A. Regev, IMAXT Consortium, G. M. Church, A. H. Marblestone, E. S. Boyden, Expansion sequencing: Spatially precise in situ transcriptomics in intact biological systems. bioRxiv 094268 [Preprint]. 15 May 2020. https://doi.org/10.1101/2020.05.13.094268.

85. C. Zhu, S. Preissl, B. Ren, Single-cell multimodal omics: The power of many. *Nat. Methods* **17**, 11–14 (2020). doi:10.1038/s41592-019-0691-5 Medline

86. E. Williams, J. Moore, S. W. Li, G. Rustici, A. Tarkowska, A. Chessel, S. Leo, B. Antal, R. K. Ferguson, U. Sarkans, A. Brazma, R. E. C. Salas, J. R. Swedlow, The Image Data Resource: A bioimage data integration and publication platform. *Nat. Methods* **14**, 775–781 (2017). doi:10.1038/nmeth.4326 Medline

87. A. C. Payne, Z. D. Chiang, P. L. Reginato, S. M. Mangiameli, zchiang/in_situ_genome_sequencing_processing: In situ genome sequencing processing code for: In situ genome sequencing resolves DNA sequence and structure in intact biological samples, Version 1.02, Zenodo (2020); https://doi.org/10.5281/zenodo.4299227.

88. A. C. Payne, Z. D. Chiang, P. L. Reginato, zchiang/in_situ_genome_sequencing_analysis: In situ genome sequencing analysis code for: In situ genome sequencing resolves DNA sequence and structure in intact biological samples, Version 1.0, Zenodo (2020); https://doi.org/10.5281/zenodo.4299125.

89. I. Solovei, A. Cavallo, L. Schermelleh, F. Jaunin, C. Scasselati, D. Cmarko, C. Cremer, S. Fakan, T. Cremer, Spatial preservation of nuclear chromatin architecture during three-dimensional fluorescence in situ hybridization (3D-FISH). *Exp. Cell Res.* **276**, 10–23 (2002). doi:10.1006/excr.2002.5513 Medline

90. C. A. Lareau, F. M. Duarte, J. G. Chew, V. K. Kartha, Z. D. Burkett, A. S. Kohlway, D. Pokholok, M. J. Aryee, F. J. Steemers, R. Lebofsky, J. D. Buenrostro, Droplet-based combinatorial indexing for massive-scale single-cell chromatin accessibility. *Nat. Biotechnol.* **37**, 916–924 (2019). doi:10.1038/s41587-019-0147-6 Medline

91. S. Picelli, Å. K. Björklund, B. Reinius, S. Sagasser, G. Winberg, R. Sandberg, Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res.* **24**, 2033–2040 (2014). doi:10.1101/gr.177881.114 Medline

92. J. H. Lee, E. R. Daugharthy, J. Scheiman, R. Kalhor, T. C. Ferrante, R. Terry, B. M. Turczyk, J. L. Yang, H. S. Lee, J. Aach, K. Zhang, G. M. Church, Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. *Nat. Protoc.* **10**, 442–458 (2015). doi:10.1038/nprot.2014.191 Medline

93. J. D. Buenrostro, B. Wu, H. Y. Chang, W. J. Greenleaf, ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* **109**, 21.29.1–21.29.9 (2015). doi:10.1002/0471142727.mb2129s109 Medline

94. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2013). doi:10.1038/nmeth.1923 Medline

95. T. Smith, A. Heger, I. Sudbery, UMI-tools: Modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* **27**, 491–499 (2017).

doi:10.1101/gr.209601.116 Medline

96. T. Massingham, N. Goldman, Error-correcting properties of the SOLiD Exact Call Chemistry. *BMC Bioinformatics* **13**, 145 (2012). doi:10.1186/1471-2105-13-145 Medline

97. S. Berg, D. Kutra, T. Kroeger, C. N. Straehle, B. X. Kausler, C. Haubold, M. Schiegg, J. Ales, T. Beier, M. Rudy, K. Eren, J. I. Cervantes, B. Xu, F. Beuttenmueller, A. Wolny, C. Zhang, U. Koethe, F. A. Hamprecht, A. Kreshuk, ilastik: Interactive machine learning for (bio)image analysis. *Nat. Methods* **16**, 1226–1232 (2019). doi:10.1038/s41592-019-0582-9 Medline

98. H. R. Ueda, A. Ertürk, K. Chung, V. Gradinaru, A. Chédotal, P. Tomancak, P. J. Keller, Tissue clearing and its applications in neuroscience. *Nat. Rev. Neurosci.* **21**, 61–79 (2020). doi:10.1038/s41583-019-0250-1 Medline

99. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010). doi:10.1093/bioinformatics/btq033 Medline

100. N. Servant, N. Varoquaux, B. R. Lajoie, E. Viara, C.-J. Chen, J.-P. Vert, E. Heard, J. Dekker, E. Barillot, HiC-Pro: An optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **16**, 259 (2015). doi:10.1186/s13059-015-0831-x Medline

101. N. C. Durand, M. S. Shamim, I. Machol, S. S. P. Rao, M. H. Huntley, E. S. Lander, E. L. Aiden, Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst.* **3**, 95–98 (2016). doi:10.1016/j.cels.2016.07.002 Medline